ORSAY

N° d'ordre : 527

THESE

PRESENTEE

A L'UNIVERSITE DE PARIS-SUD

CENTRE D'ORSAY

POUR OBTENIR

LE TITRE DE DOCTEUR INGENIEUR

par

Michel INVERNIZZI

SURVEILLANCE D'UN COEUR DE REACTEUR NUCLEAIRE PAR UNE METHODE DE RECONNAISSANCE DE FORMES

Soutenue le 17 mai 1982, devant la Commission d'Examen

MM.	PICINBONO	Président
	GUEGUEN	
	POTTIER	Examinateurs
	DUPEYRAT	
	LIVOLANT	

ORSAY
N° d'ordre : 527

THESE

PRESENTEE

A L'UNIVERSITE DE PARIS-SUD
CENTRE D'ORSAY

POUR OBTENIR

LE TITRE DE DOCTEUR INGENIEUR

par

Michel INVERNIZZI

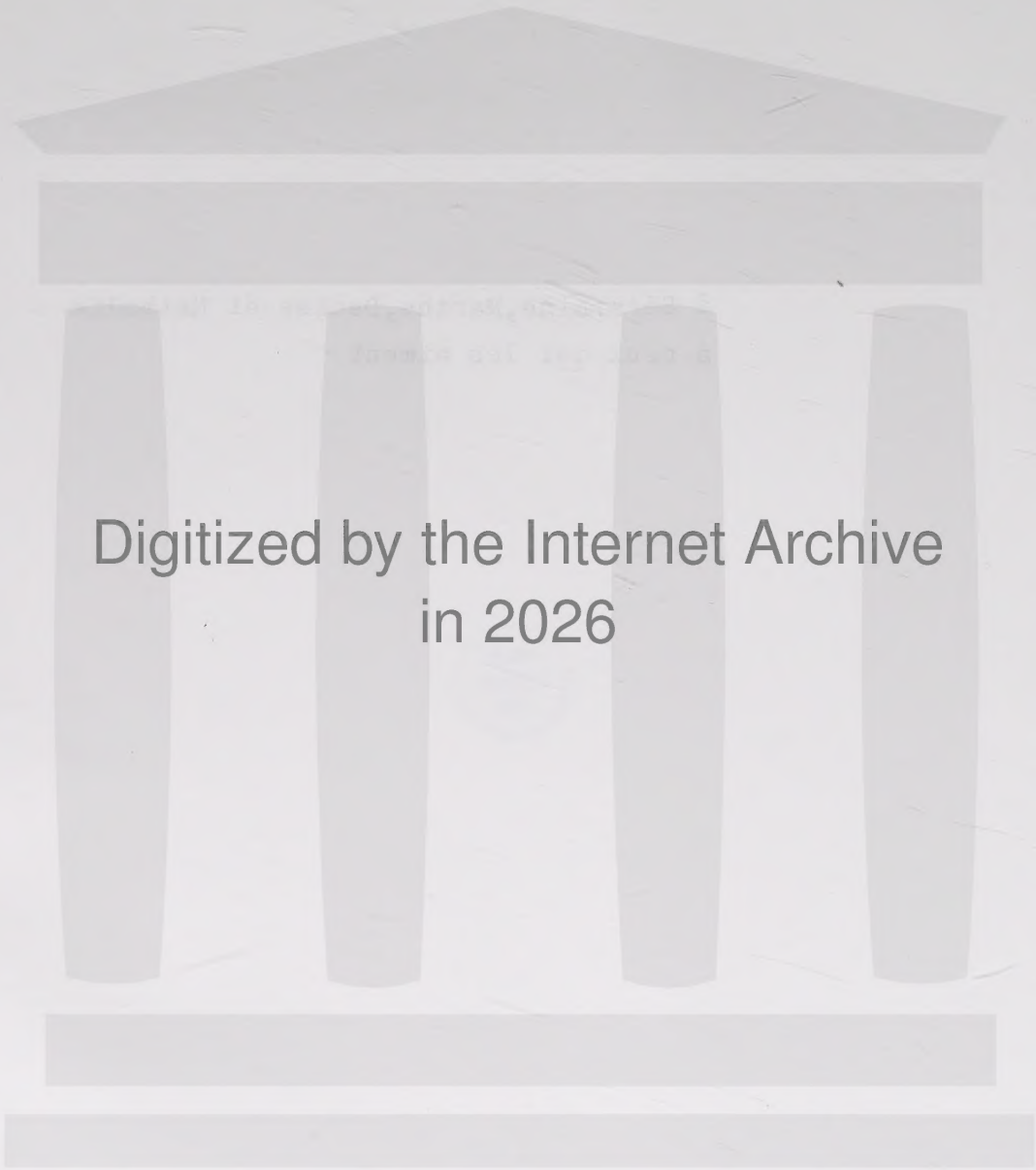
SURVEILLANCE D'UN COEUR DE REACTEUR NUCLEAIRE PAR UNE METHODE DE RECONNAISSANCE DE FORMES

Soutenue le 17 mai 1982, devant la Commission d'Examen

MM.	PICINBONO	Président
	GUEGUEN	
	POTTIER	Examinateurs
	DUPEYRAT	
	LIVOLANT	

UNIV.
LEIDEN
BIBL.

à Séraphine, Marthe, Denise et Nathalie
à ceux qui les aiment


Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41543523_0026

AVANT-PROPOS

Je désire remercier Monsieur le professeur PICINBONO de l'intérêt qu'il a bien voulu porter à mon travail.

Je le remercie d'avoir accepté de présider le jury de cette thèse.

Que Messieurs GUEGUEN, LEROUX, PRADO, MICLET du département "Systèmes et Communications" de l'Ecole Nationale Supérieure des Télécommunications, qui m'ont initié avec sympathie aux techniques de reconnaissance de formes, trouvent ici l'expression de ma gratitude.

Je remercie Messieurs GUEGUEN, POTTIER, LIVOLANT, DUPEYRAT pour l'honneur qu'ils me font de participer à ce jury.

Cette étude a été réalisée au Département d'Electronique et d' Instrumentation Nucléaire du Centre d'Etudes Nucléaire de Saclay grâce à l'accueil et à la bienveillance de Monsieur WEILL chef du département et de Monsieur JOVER, chef du service des applications industrielles.

Je remercie Monsieur POUJOL ainsi que tous les membres de son groupe pour l'aide amicale qu'ils m'ont témoignée.

SOMMAIRE

INTRODUCTION

1 - RECONNAISSANCE STATISTIQUE DES FORMES

1.1. - Présentation
1.2. - Chaîne de reconnaissance

1.2.1. - Les capteurs
1.2.2. - Le prétraitement
1.2.3. - L'apprentissage
1.2.4. - La décision
1.2.5. - Les réactions

1.3. - Cadre de l'étude

2 - RECONNAISSANCE DE FORMES ET SURVEILLANCE DE PROCESSUS

2.1. - Introduction
2.2. - Hypothèse de normalité constante
2.3. - Hypothèse de normalité faiblement évolutive
2.4. - Surveillance et diagnostic
2.5. - Champ d'application

3 - SURVEILLANCE D'UN COEUR DE REACTEUR NUCLEAIRE PAR UNE METHODE DE RECONNAISSANCE DE FORMES

3.1. - Introduction

3.1.1. - Bruit neutronique et comportement vibratoire du coeur
3.1.2. - But de l'étude
3.1.3. - Principe retenu
3.1.4. - Notations

3.2. - Les données
3.3. - Le prétraitement

3.3.1. - Prétraitement à priori
3.3.2. - Prétraitement à postériori
3.3.3. - Commentaires sur la réduction

3.4. - Classification automatique

3.4.1. - Classification automatique en fréquence
3.4.2. - Distance inter-résonances, distance inter-formes
3.4.3. - Classification automatique en distance
3.4.4. - Algorithme des k moyennes
3.4.5. - Commentaires sur l'apprentissage

3.5. - La décision

3.5.1. - Estimation des distributions des classes
3.5.2. - Décision baysienne
3.5.3. - Décision d'anormalité

3.6. - Application informatique

3.6.1. - Le programme "APREN"
3.6.2. - Le programme "RECON"
3.6.3. - Visualisation des classes

3.7. - Résultats expérimentaux

3.7.1. - Paramètres utilisés
3.7.2. - Résultats généraux
3.7.3. - Probabilité de fausse-alarme
3.7.4. - Détection d'une fréquence parasite
3.7.5. - Bruit blanc parasite

3.8. - Fonctionnement temps réel

3.8.1. - Apprentissage hors ligne
3.8.2. - Surveillance en ligne

4 - ADAPTATIVITE DE LA RECONNAISSANCE DE FORMES

4.1. - Introduction
4.2. - Approches classiques
4.3. - Reconnaissance adaptative
4.4. - Reconnaissance adaptative par modélisations sphériques

4.4.1. - Présentation de la méthode
4.4.2. - La phase d'apprentissage
4.4.3. - Résultats expérimentaux
4.4.4. - La phase de reconnaissance adaptative
4.4.5. - Résultats expérimentaux
4.4.6. - Perspectives sur données nucléaires

5 - CONCLUSION

INTRODUCTION

La reconnaissance des formes est l'une des disciplines ayant connu l'un des plus grands développements lors de ces vingt dernières années. Ses applications, d'abord limitées à quelques domaines particuliers (zoologie, botanique, ...) sont de plus en plus nombreuses. La physique, la chimie, la linguistique, la psycologie, la biologie, la médecine, la météorologie, la zoologie, la géologie, ... et bien d'autres disciplines utilisent déjà couramment des techniques de reconnaissance. Cette diffusion est d'autant plus remarquable que, malgré un effort lors de ces cinq dernières années, le nombre d'ouvrages de base et de synthèse est encore très réduit et que le mot "forme" a un sens courant trop restrictif.

Cette étude propose l'application de la reconnaissance de formes au problème de la surveillance et de la détection précoce d'incidents de fonctionnement de processus complexes.

Le chapitre 1 présente sommairement la reconnaissance de formes.

Au chapitre 2 nous montrerons comment la reconnaissance de formes peut être utilisée pour surveiller un processus et soulèverons le problème de l'évolution des critères de jugement.

Le chapitre 3 expose un algorithme de reconnaissance de formes spécialement étudié pour la surveillance du coeur de la centrale nucléaire de Bugey II.

Au chapitre 4 nous essayerons d'apporter quelques éléments de réponse au problème de l'évolution des critères et proposerons un algorithme original de reconnaissance adaptative basé sur une modélisation sphérique des classes.

1 - LA RECONNAISSANCE STATISTIQUE DES FORMES

1.1. - Présentation

La reconnaissance des formes est une discipline n'ayant pris sa vraie signification qu'avec le développement des ordinateurs. Elle traite de l'automatisation informatique de taches de "perceptions artificielles" réalisées habituellement par le cerveau humain. Un exemple souvent donné est l'identification automatique de caractères manuscrits. Mais les techniques de reconnaissance sont également utilisées pour identifier les différentes phases du sommeil dans un électro-encéphalogramme ou pour effectuer des diagnostics médicaux par ordinateur.

La seule notion de "forme" pose des problèmes : le mot anglais "Pattern" est moins restrictif il sous-entend une notion de modèle type, de patron, de moule qui disparait en français. Notre propos n'est pas de diserter sur la notion de forme. Contentons-nous de signaler les travaux de Dydack and Segal [1] et Pavel [2] établissant une théorie de la forme.

Nous admettrons comme "forme" un état du monde extérieur à l'ordinateur sur lequel celui-ci possède des informations.

1.2. - Chaîne de reconnaissance

Le schéma général d'une chaîne de reconnaissance est le suivant :

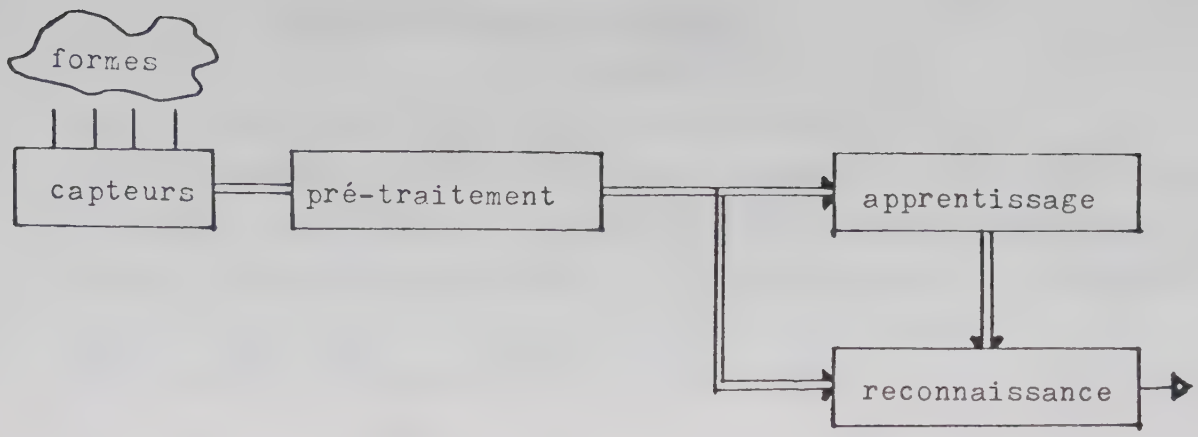


Fig 1: Schéma d'une chaine de reconnaissance

Nous pouvons distinguer dans ce schéma 4 modules.

1.2.1. - Les capteurs

Nous appelons "capteurs" l'interface entre les formes à étudier et l'ordinateur. C'est à dire le canal transmettant les informations. Le premier problème de la reconnaissance des formes est le choix des capteurs. Ceux-ci doivent transmettre des informations "utiles" relativement au problème à résoudre et aux méthodes allant être utilisées. La détermination à priori de cette utilité reste du ressort de l'expérimentateur. Lorsque celui-ci ne possède que peu d'informations à priori il est souvent amené à utiliser un nombre surabondant de capteurs.

1.2.2. - Le prétraitement

Le prétraitement comprend deux parties. La première ne dépend pas des données et consiste à mettre en forme les variables fournies par les capteurs (exemple : contrôle automatique de gain, filtrage, FFT, ...). La deuxième fait l'objet d'un apprentissage et a trois rôles fondamentaux.

1.2.2.1. - L'extraction de paramètres de reconnaissance

Les capteurs fournissent une masse de données souvent énorme. L'information "utile"* est diffuse parmi toutes les variables. Extraire des paramètres de reconnaissance c'est effectuer une transformation de l'espace de ces variables dans un espace de dimension moindre de telle façon que l'information utile soit concentrée sur un nombre de paramètres aussi faible que possible.

1.2.2.2. - La sélection des paramètres de reconnaissance

Elle consiste à déterminer parmi les paramètres définis à l'étape précédente ceux qui sont les plus "utiles".

1.2.2.3. - L'ordonnancement des paramètres de reconnaissance

Certaines chaînes de reconnaissance sont séquentielles : elles considérent les paramètres de reconnaissance successivement et prennent à chaque instant une décision entre les alternatives suivantes :

- Effectuer la classification sur la base des renseignements déjà acquis.
- Prendre en compte un nouveau paramètre pour afiner la décision.

Il se pose alors le problème de l'ordre dans lequel les différents paramètres doivent être présentés.

---

* La notion d'utilité peut être abordée par la théorie de l'information, l'analyse de données, l'optimisation de critères de représentativité et/ou de séparabilité des classes... Elle se conçoit relativement au taux de bonne reconnaissance.

1.2.3. - L'apprentissage

Le module d'apprentissage a pour rôle de définir les règles de décisions allant être utilisées par la suite. Il y a souvent également apprentissage d'une partie du pré-traitement. L'apprentissage est basé sur l'observation d'un certain nombre de formes dites d'initialisation ou d'apprentissage. Ces formes sont captées puis pré-traitées. Elles sont donc transformées en un certain nombre de jeux de paramètres que l'on appelle base d'apprentissage. Deux types de techniques peuvent alors être utilisés selon que :

- On connait le nombre de classes possibles et l'on sait à quelle classe chaque forme d'apprentissage appartient. On dit qu'il y a apprentissage avec professeur.
- On connait ou non le nombre de classes mais l'on ignore à quelle classe appartient chaque forme d'apprentissage. On dit qu'il y a apprentissage sans professeur. Les algorithmes devront dans ce cas déterminer le nombre de classes si ce nombre est inconnu et partitionner la base d'apprentissage sur ces classes. Ces techniques relèvent de la classification automatique. Lorsque cela est possible, les classes générées sont ensuite qualifiées par l'expérimentateur qui cherche ainsi à leur donner une signification physique.

1.2.4. - La décision

Elle consiste, postérieurement à l'apprentissage, à affecter une observation inconnue à une classe prédéterminée. C'est la phase de reconnaissance proprement dite.

1.2.5. - Les réactions

La présentation faite précédemment d'un schéma de reconnaissance est très sommaire : on ne peut pas traiter séquentiellement les différents modules et trouver indépendamment une optimisation du choix des capteurs, une méthode de paramètrisation, un algorithme d'apprentissage et une technique de décision. La justesse des choix ne peut être vérifiée qu'en fin de traitement et postérieurement à un grand nombre d'expériences. Cela implique en général une grande complexité des algorithmes d'apprentissage dont les caractéristiques doivent être adaptatives. La présence de ces contre-réactions pose des problèmes d'optimisation qui sont souvent insolubles mathématiquement sauf sous hypothèses abusivement restrictives. Les meilleures techniques de reconnaissance actuellement utilisées sont le résultat de compromis entre critères d'expérience et critères formels mathématiques. La reconnaissance des formes est encore un domaine ouvert à l'heuristique.

1.3. - Cadre de l'étude

En raison de l'application désirée, l'étude entreprise ci-après entre dans le cadre de la reconnaissance statistique des formes et est restreinte aux cas suivants :

- Les capteurs fournissent un secteur $\vec{x}$ de $\mathbb{R}^n$
- Le prétraitement est une transformation de $\mathbb{R}^n$ dans $\mathbb{R}^P$ avec $P \leq n$.
- L'apprentissage décisionnel génére un partitionnement de $\mathbb{R}^P$.
- La décision est une application de $\mathbb{R}^P$ dans l'ensemble des indices de classes. (On ne considérera pas le cas de la décision séquentielle).

Nous ne nous poserons pas le problème du choix des capteurs. Ceux qui seront utilisés lors de nos applications seront présentés au chapitre 3.

Il est hors de propos de vouloir présenter ne serait-ce qu'un aperçu des algorithmes de reconnaissance de formes existant. Il y en a autant que de problèmes particuliers ayant été traités. En effet, un tel algorithme est composé de trois parties principales (extraction/sélection de paramètres, apprentissage, décision) et d'une méthode assurant le lien et la compatibilité entre elles. Nous présentons en annexe à la présente étude une liste de références bibliographiques concernant chacune de ces trois parties. L'étude présentée en référence [31] est un recueil des méthodes et des algorithmes utilisés en reconnaissance de formes.

2 - RECONNAISSANCE DE FORMES ET SURVEILLANCE DE PROCESSUS

2.1. - Introduction

Le processus à surveiller est un système dynamique. Dans le domaine de l'analyse du réel un système dynamique est constitué par les liaisons entre grandeurs représentatives observées d'un phénomène, ces grandeurs évoluant en fonction du temps et de causes externes. La surveillance d'un tel système est souvent effectuée par des procédés de seuillage empiriques et plus rarement par le moyen de modèlisations. Modèliser un système dynamique réel consiste à lui faire correspondre un système d'équations formalisant mathématiquement les correspondances. Dans une phase dite d'apprentissage, un modèle mathématique est déterminé. Par la suite, le processus est contrôlé en comparant les sorties réelles et les sorties fournies par le modèle fonctionnant dans les mêmes conditions. Lorsque ces sorties différent la décision d'anormalité est prise.

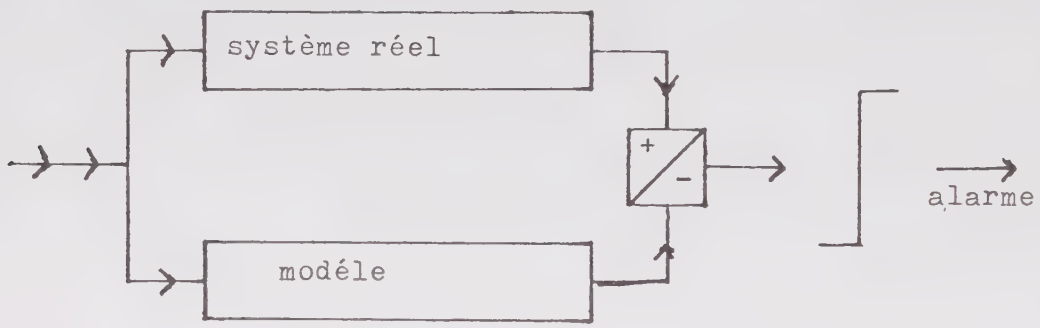
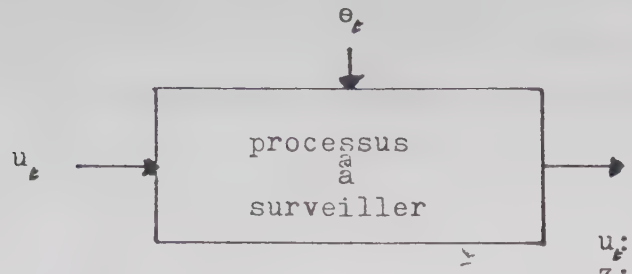


Fig 2: Surveillance par modélisation

L'utilisation des techniques de reconnaissance de formes pour résoudre des problèmes de surveillance ou de diagnostic est plus récente. Les références [3] à [20] illustrent quelques unes des applications réalisées dans le domaine de l'aéronautique et de la médecine. Le diagnostic ou la surveillance automatique consistent à reconnaitre des causes ou des états de panne à partir des symptômes observés et des antécédents. Il y a, à la limite, classification par association des symptômes à une maladie ou à une cause unique. Si ces symptômes peuvent être mis sous la forme d'un vecteur mesurable, cette observation constitue une "forme".

Une surveillance automatique consiste à reconnaître automatiquement la "forme-symptôme" dans l'ensemble des formes possibles classifiées par "classes-modes de fonctionnement".

La présente étude ne suppose pas la connaissance d'un modèle. Le système y est considéré comme une boîte noire :



u_t: entrées connues
z_t: observations possibles
θ_t: processus aléatoire

Fig 3: Surveillance statistique

A chaque instant, certains modes de fonctionnement sont considérés comme normaux, d'autres comme anormaux. Surveiller le processus consiste à reconnaitre en ligne ce caractère.

Notons :

Ω_n : L'ensemble des valeurs possibles de θ au temps t_n.

Ω_n^N : L'ensemble des valeurs de θ considérées comme normales au temps t_n.

$\Omega_n^A = \Omega_n - \Omega_n^N$ l'ensemble des valeurs de θ considérées comme anormales au temps t_n.

χ_n = L'ensemble des formes possibles au temps t_n

Dans le cadre de cette étude, nous appellerons "forme" une réalisation du secteur $x_n = (z_n', u_n .. u_{n-p})$ et supposerons qu'une observation effectuée au temps t_n est indépendante des entrées antérieures au temps t_{n-p}.

Nous distinguons ci-après deux cas relativement à la façon dont évolue la notion de normalité au cours du temps.

2.2. - Hypothèse de normalité constante

La notion de normalité est indépendante du temps

$\forall n \quad \Omega_n = \Omega \quad \Omega_n^N = \Omega^N \quad \Omega_n^A = \Omega^A \quad \chi_n = \chi$

$\Omega^A \cap \Omega^N = \phi$

$\Omega^A \cup \Omega^N = \Omega$

Cela signifie qu'une forme porte de façon indépendante du temps le critère normal ou anormal. Dans ce cas le problème peut être supposé ergodique et ramené à celui classique de reconnaissance à deux classes qui consiste à établir une règle de décision δ de l'ensemble des formes possibles χ dans l'ensemble à deux éléments $A = \{\Omega^N, \Omega^A\}$. La base d'apprentissage B sera constituée d'épreuves successives de $\vec{x}$. $B = \{\vec{x}_{t_1} \ .... \ \vec{x}_{t_N}\}$

Exemple 1 :

Considérons le système dynamique suivant :

$$y_k = \frac{1}{2}\theta_1 u_k + \theta_2 u_k^2 + W$$

avec $u \in [0.6, 1]$.

θ_1 et θ_2 variables aléatoires à valeurs dans $\{0,1\}$

W un bruit de loi $N(0, \sigma^2)$.

Sous l'hypothèse de normalité :

$\Omega_N = \{(0, 0, w), (0, 1, w), (1, 0, w) \ \forall w\}$

$\Omega_A = \{(1, 1, w) \ \forall w\}$

Supposons qu'après un certain nombre d'expériences la base d'apprentissage ci-dessous ait été formée.

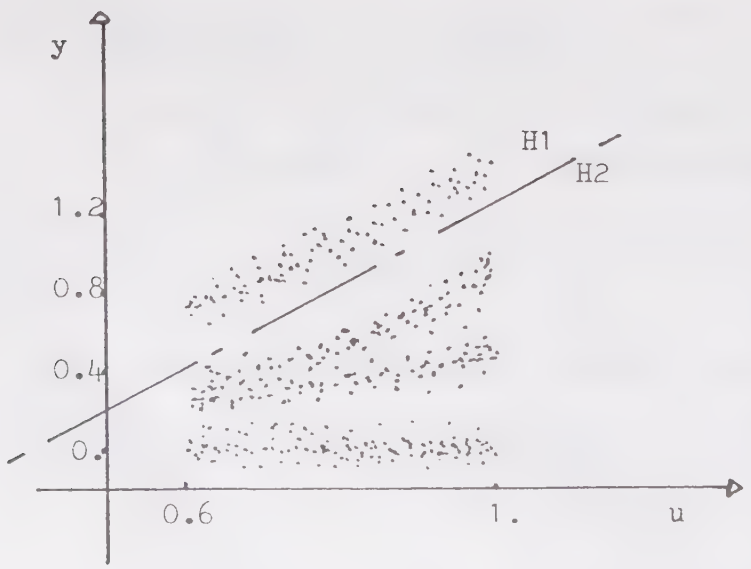


Fig 4: Base d'apprentissage dans le plan (u,y)

Un algorithme de reconnaissance de formes pourrait par exemple générer la règle de décision linéaire suivante :

$x \in H_1 \longrightarrow \theta$ anormal $\qquad x = (u, y)$

$x \in H_2 \longrightarrow \theta$ normal.

Exemple 2 :

$$y_k = y_{k-1} + u_k + \theta_1 + N(0, \sigma^2)$$

$$u_k \in \{-1, 0, +1\}$$

θ_1 variable aléatoire à valeurs dans $\{-1, 0 + 1\}$

$\Omega_N = \{0\} \qquad \Omega_A = \{-1, +1\}$

Ce système n'est pas à mémoire finie relativement aux entrées mais si nous considérons l'observation : $z_k = y_k - y_{k-1}$ nous obtenons :

$$z_k = u_k + \theta_1 + N(0, \sigma^2)$$

Une base d'apprentissage pourrait être :

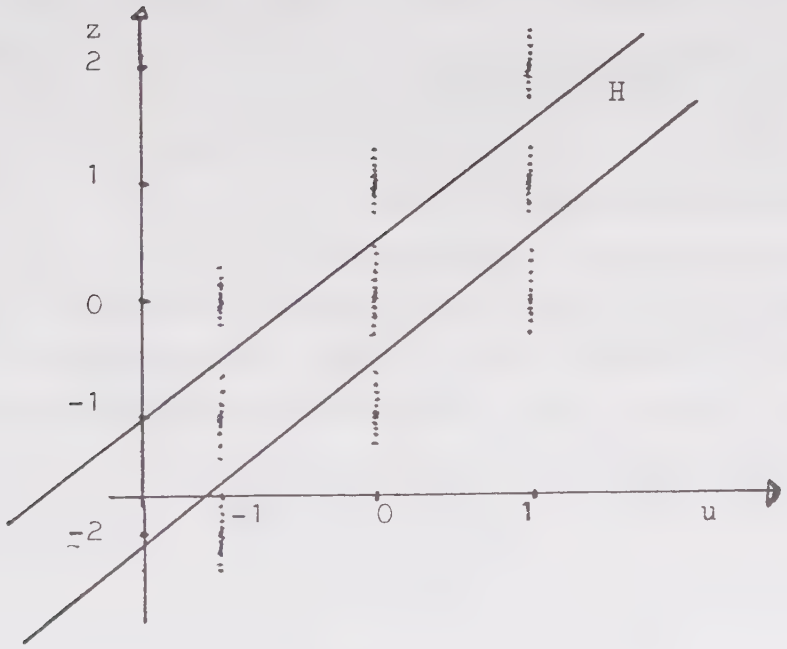


Fig 5: Base d'apprentissage dans le plan (u,z)

Une règle de décision pouvant être générée par un algorithme de reconnaissance des formes est :

$z \in H \longrightarrow \theta$ normal

$z \notin H \longrightarrow \theta$ anormal

2.3. - Hyptohèse de normalité faiblement évolutive

Il n'est pas toujours justifié de supposer la notion de normalité constante. Un phénomène considéré comme anormal dans un processus venant d'être initialisé peut être admis dans ce même processus ayant vieilli. C'est le cas par exemple d'un réacteur nucléaire dont le comportement varie en fonction de l'âge du combustible. Nous appellerons "cycle" la durée de vie d'un tel processus. L'ergodicité n'étant plus vérifiée, chaque instant pose son propre problème de reconnaissance.

A chaque instant t_n, déterminer une fonction de décision de χ_n dans $A_n = \{\Omega_n^N, \Omega_n^A\}$ basée sur l'observation de la base d'apprentissage $B_n = \{\vec{x}_n$ provenant de cycles différents$\}$.

Bien que en théorie faisable, cette approche serait extrémement lourde. D'autre part, d'un cycle à l'autre les conditions de fonctionnement, de vieillissement et de mesure peuvent différer. Il faudrait alors à chaque instant t_A considérer tout le passé du processus !

Nous admettrons que l'évolution de la notion de normalité est lente vis à vis de la fréquence d'acquisition des observations de telle façon que N épreuves successives puissent être considérées comme provenant d'une même notion de normalité. La base d'apprentissage au temps t_k peut alors être formée des épreuves $\vec{x}_{k-1}$ ... $\vec{x}_{k-N}$ provenant éventuellement de plusieurs cycles. Chaque terme pouvant voir son importance pondérée par son antériorité.

2.4. - Surveillance et diagnostic

Les paragraphes précédents montrent comment il est parfois possible de ramener le problème de la surveillance d'un processus complexe à celui d'une reconnaissance de formes à deux classes (Normal, Anormal). Cette approche peut toutefois être affinée et nuancée suivant le but poursuivi et les données disponibles. Au chapitre 1, nous avons présenté deux types d'apprentissage (supervisé ou non) suivant que les formes d'apprentissage aient été préalablement identifiées ou non. En fait, il est possible d'avoir des bases de données partiellement supervisées. Par exemple, une base peut être supervisée pour les formes normales (pour chaque forme normale on sait qu'elle provient d'un certain type de fonctionnement) et non supervisée pour les formes anormales. On peut alors s'intéresser au problème de diagnostic c'est à dire l'identification des types de fonctionnement dégradés. Dans d'autres cas, l'on ne disposera que d'une base d'apprentissage de formes normales. L'anormalité devra alors être définie par opposition. Nous présentons ci-dessous quelques cas possibles d'utilisation de la reconnaissance de formes.

1) - La base d'apprentissage ne contient que des formes normales

. La normalité est considérée globalement.

L'anormalité est alors définie par un test statistique. Il n'y a pas à proprement parler reconnaissance mais apprentissage d'une mesure d'écart. C'est l'approche suivie par Al Mouhamed [21].

. On cherche à analyser les différentes façons pour une forme d'être normale,

Le problème peut alors être posé comme celui d'une reconnaissance avec apprentissage non supervisé :

Une méthode de classification automatique opérant sur la base d'apprentissage des formes normales peut mettre en évidence un certain nombre de classes que l'on essayera par la suite d'identifier à des types de fonctionnement ou a des phénomènes physiques. La décision d'anormalité est alors prise lorsque la classification d'une nouvelle forme parmi ces classes normales s'avère impossible ou peu vraisemblable. Cette approche est celle suivie par l'application du chapitre 3.

. Dans le cas où une classification des formes normales existe déjà avant l'apprentissage, l'approche proposée ci-dessus reste valable et s'en trouve simplifié .

2) - La base d'apprentissage contient des formes normales et anormales :

. Mais l'on ne connnait pas le caractère normal-anormal de chaque forme. Ce problème peut être vu comme une reconnaissance à deux classes avec apprentissage non supervisé.

. Si l'on connait le caractére de chaque forme. L'apprentissage sera supervisé sur ce caractère. Il sera alors possible d'affiner l'analyse en cherchant une classification parmi les formes anormales (diagnostic) et/ou parmi les formes normales (contrôle).

. Si l'une ou l'autre de ces deux classifications est connue le problème s'en trouve simplifiée.

2.5. - Champ d'application

L'approche que nous proposons présente un certain nombre de limitations ; l'un des problèmes clefs de la reconnaissance de formes est la réduction du nombre de paramètres représentant chaque forme. Le fait de considérer comme forme les vecteurs $x_k = (z'_k, u'_k \ldots u'_{k-p})$ impose que le nombre d'entrées qu'il est nécessaire de prendre en compte soit faible, c'est à dire que l'observation z_k doit être choisie la plus indépendante possible des entrées.

Les nuages de formes correspondant aux différents modes de fonctionnements sont sous tendus par la variété géométrique modélisant les liaisons observations/entrées du système réel. Ces liaisons sont donc néfastes aux hypothèses de distributions gaussiennes généralement émises.

Malgré ces considérations négatives un certain nombre de processus peut être surveillé par cette approche. C'est le cas en particulier des processus autorégulés admettant peu ou pas d'entrées. Un réacteur nucléaire en fonctionnement stationnaire peut être considéré comme un tel processus.

3 - SURVEILLANCE D'UN COEUR DE REACTEUR NUCLEAIRE PAR UNE METHODE DE RECONNAISSANCE DE FORMES

3.1. - Introduction

Plusieurs méthodes de surveillance d'un réacteur nucléaire basées sur une approche par la reconnaissance des formes ont été proposées. Nous pouvons citer les travaux de CONZALEZ et HOWINGTON [22-25] sur le réacteur à haut flux isotopique du laboratoire Oak Ridge (Tennesse), et ceux de DUBUISSON-LAVISON [26-28] de l'Université de technologie de Compiègne portant sur la surveillance d'un réacteur à neutrons rapides.

La méthode de surveillance proposée ci-après a été établie et expérimentée sur la base d'un ensemble de mesures réalisées sur le réacteur nucléaire de Bugey II. Ce réacteur est du type PWR (pressurized-water-reactor). Il comporte deux circuits d'eau séparés. Un circuit primaire refroidit le coeur et joue un rôle de modérateur. Ce circuit céde son énergie via un échangeur au circuit secondaire chargé d'alimenter en vapeur la turbine. Le combustible utilisé est de l'uranium enrichi. Le schéma de principe d'une tranche PWR est donné figure 6. Sur un tel réacteur de nombreux capteurs sont installés. Il s'agit essentiellement de chambres neutroniques, d'accéléromètres de capteurs de pression de débit, de niveau ou de température. La figure 7 présente une liste de quelques un de ces capteurs et la figure 8 montre leur implantation.

Une méthode de surveillance statistique suppose la réalisation d'un grand nombre de mesures simultanées sur les capteurs et échelonnées tout au long d'un cycle de combustible. Les seuls capteurs ayant fait l'objet d'une telle campagne de mesure sont les chambres neutroniques du coeur. Le problème du choix des capteurs qui fait en principe partie de la détermination d'une méthode de reconnaissance de formes ne se pose donc pas. Ce qui soulève une question primordiale : les signaux issus des chambres neutroniques sont-ils porteurs ou non d'informations pertinentes relativement au problème de surveillance ?

3.1.1. - Bruit neutronique et comportement vibratoire du coeur

Le coeur de la centrale nucléaire de Bugey II est formé de 157 assemblages combustibles de section carrée . La figure 9 montre en coupe horizontale la position des quatre chambres. Chacune de ces chambres possède deux sections superposées (haute et basse). Leur rôle essentiel est de mesurer le flux neutronique émis et par là la puissance "neutronique" délivrée par le coeur. Schématiquement elles se composent d'une enceinte remplie de gaz dans laquelle plongent deux électrodes entre lesquelles on établit un champ électrique. Lorsqu'une particule ionisante traverse le volume gazeux, elle crée un certain nombre de paires d'ions qui se dirigent vers l'une ou l'autre des électrodes suivant le signe de leur charge. Un certain nombre de charges est recueilli par les électrodes et constitue le signal neutronique. Les particules ionisantes étant crées par des collisions entre neutrons et atomes du gaz, ce signal mesure le flux neutronique. Les informations en général extraites de ce signal se résument aux fluctuations de sa valeur moyenne. Nous nous intéressons ici à son côté bruit . En effet, il est apparu [29-30] qu'une analyse fréquentielle du bruit neutronique permettait de mettre en évidence certains comportements vibratoires du coeur. La figure 10 montre une densité spectrale de puissance obtenue par analyse de Fourier du bruit neutronique.

L'étude des cohérences entre les différentes voies a permi de déterminer les modes de ces vibrations.

Le tableau de la figure 11 résume les résultats obtenus.

On considère que les caractéristiques des résonances comprises entre 0 et 30 hz des spectres neutroniques sont des paramètres importants à surveiller.

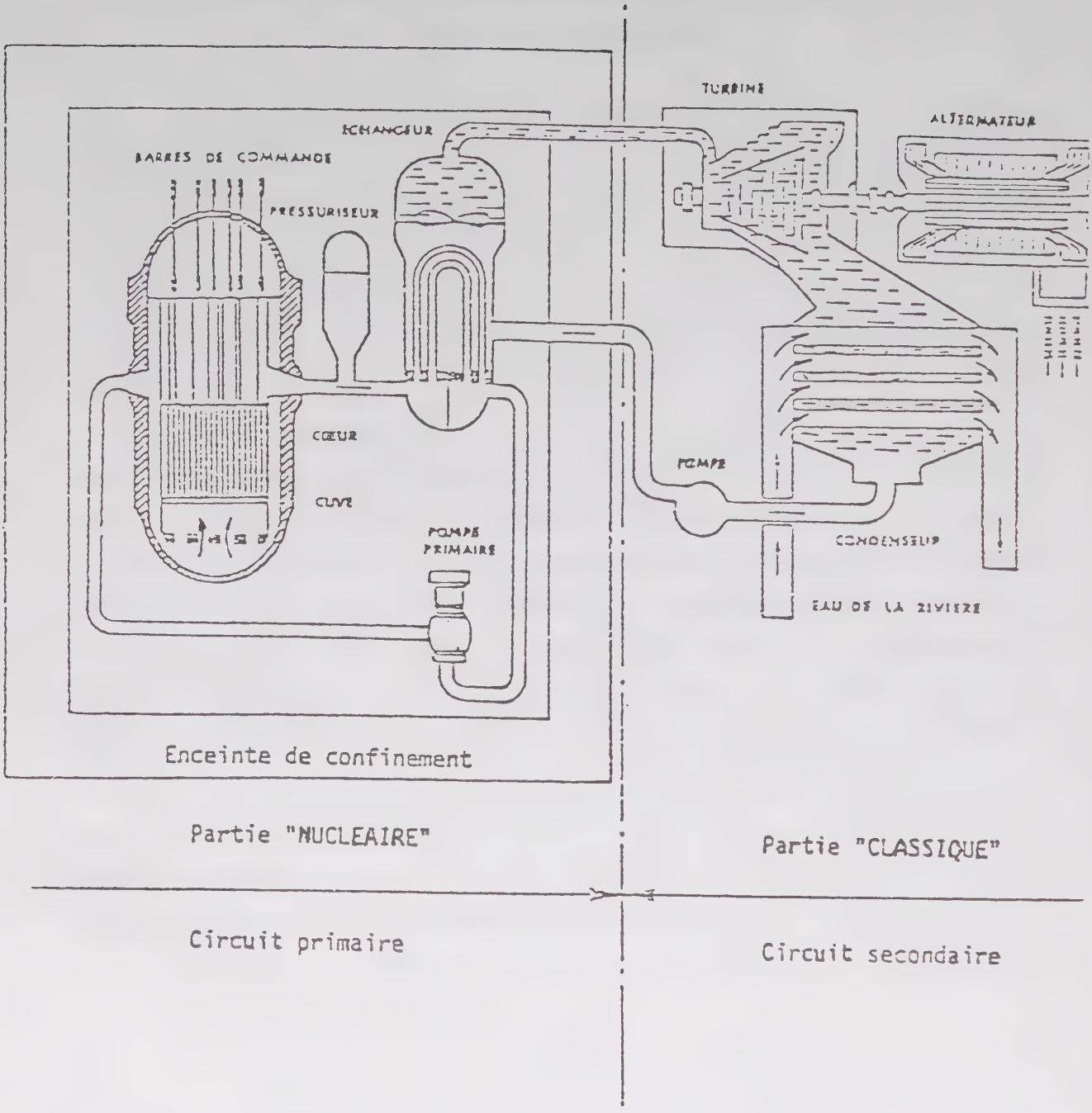


Fig.6 - Schéma de principe d'une tranche nucléaire PWR 900 MWe.

1B	Chambre neutronique basse	n° 1
2B	"	n° 2
3B	"	n° 3
4B	"	n° 4
1 H	Chambre neutronique haute	n° 1
2 H	"	n° 2
3 H	"	n° 3
4 H	"	n° 4
AC1	Accéléromètre de cuve	fond de cuve horizontal
AC2	"	"
AC3	"	couvercle horizontal tangentiel
AC4	"	couvercle vertical
AC5	"	couvercle horizontal radial
AC6	"	couvercle vertical
AC7	"	couvercle horizontal tangentiel
AP1	Accéléromètre corps de pompe primaire	n° 1
AP2	"	n° 2
AP3	"	n° 3
AG1	Accéléromètre générateur de vapeur	n° 1
AG2	"	n° 2
AG3	"	n° 3
PP1	Capteur de pression tuyauterie primaire	n° 1 refoulement
PP2	"	n° 2 "
PP3	"	n° 3 "
PP4	Capteur de pression branche chaude RRA	

Fig 7: Nomenclature des capteurs et repérage

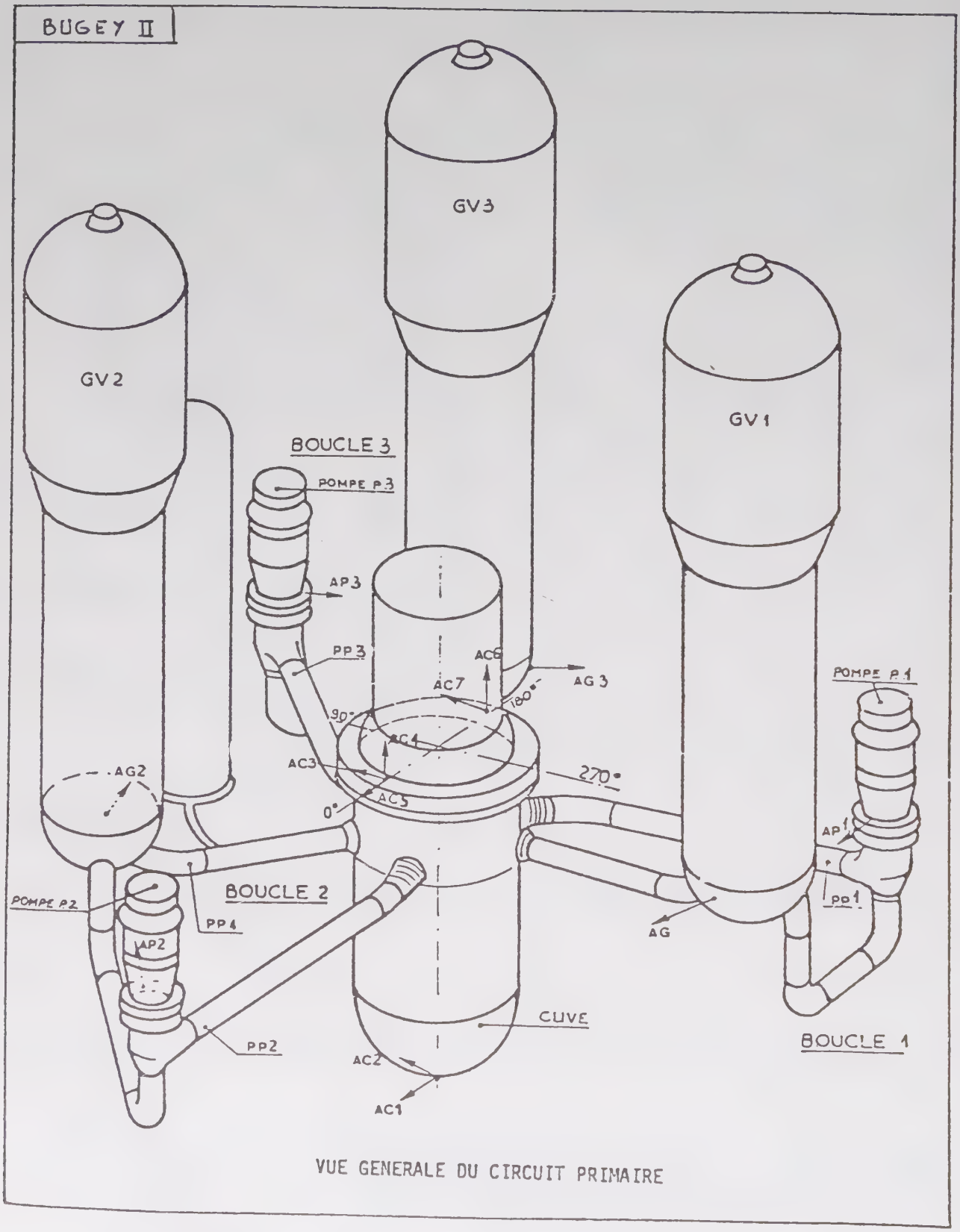
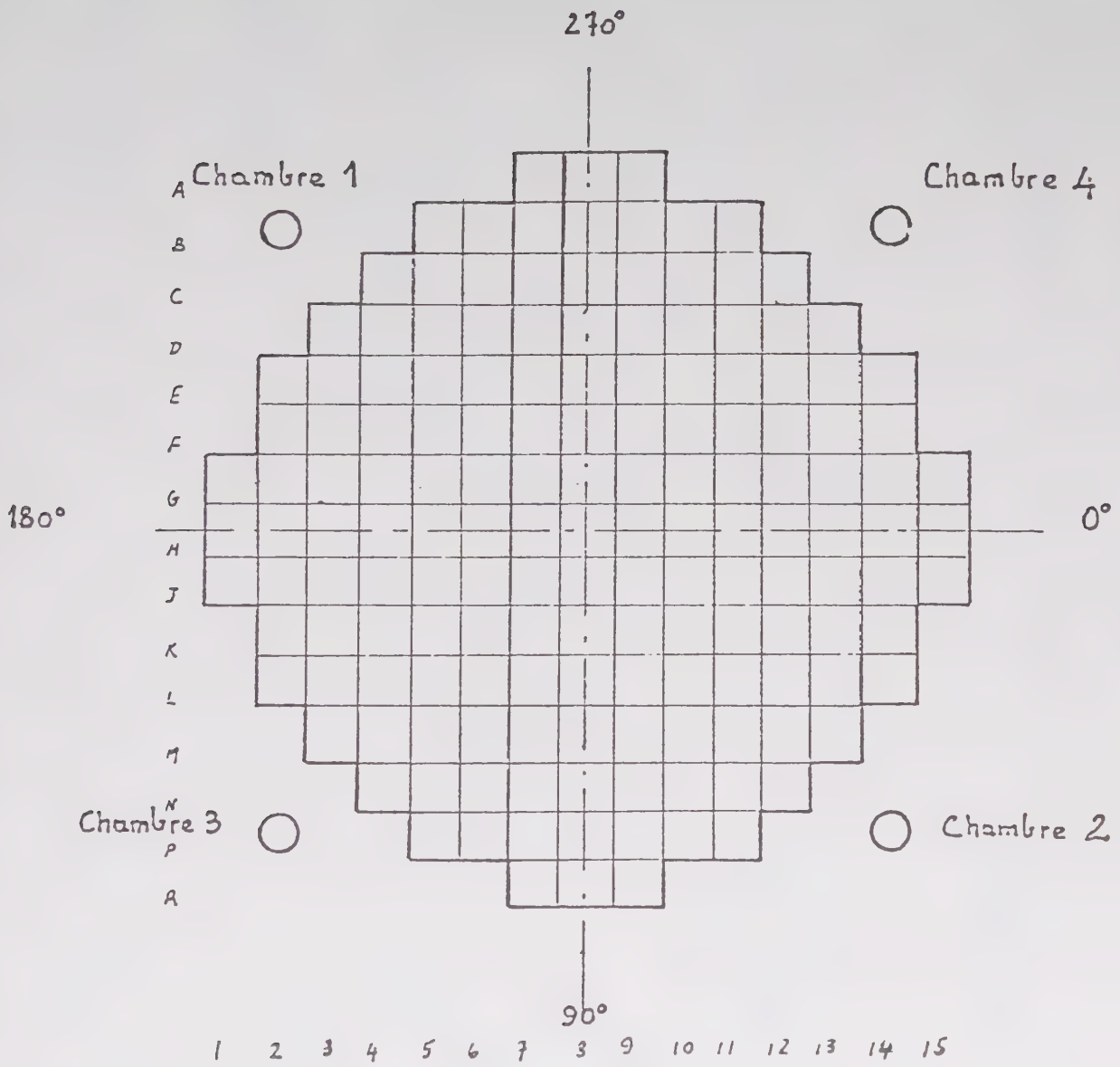


Fig 8: Schéma d'implantation des capteurs

Fig 9: Position des Chambres Neutroniques

BUGEY II



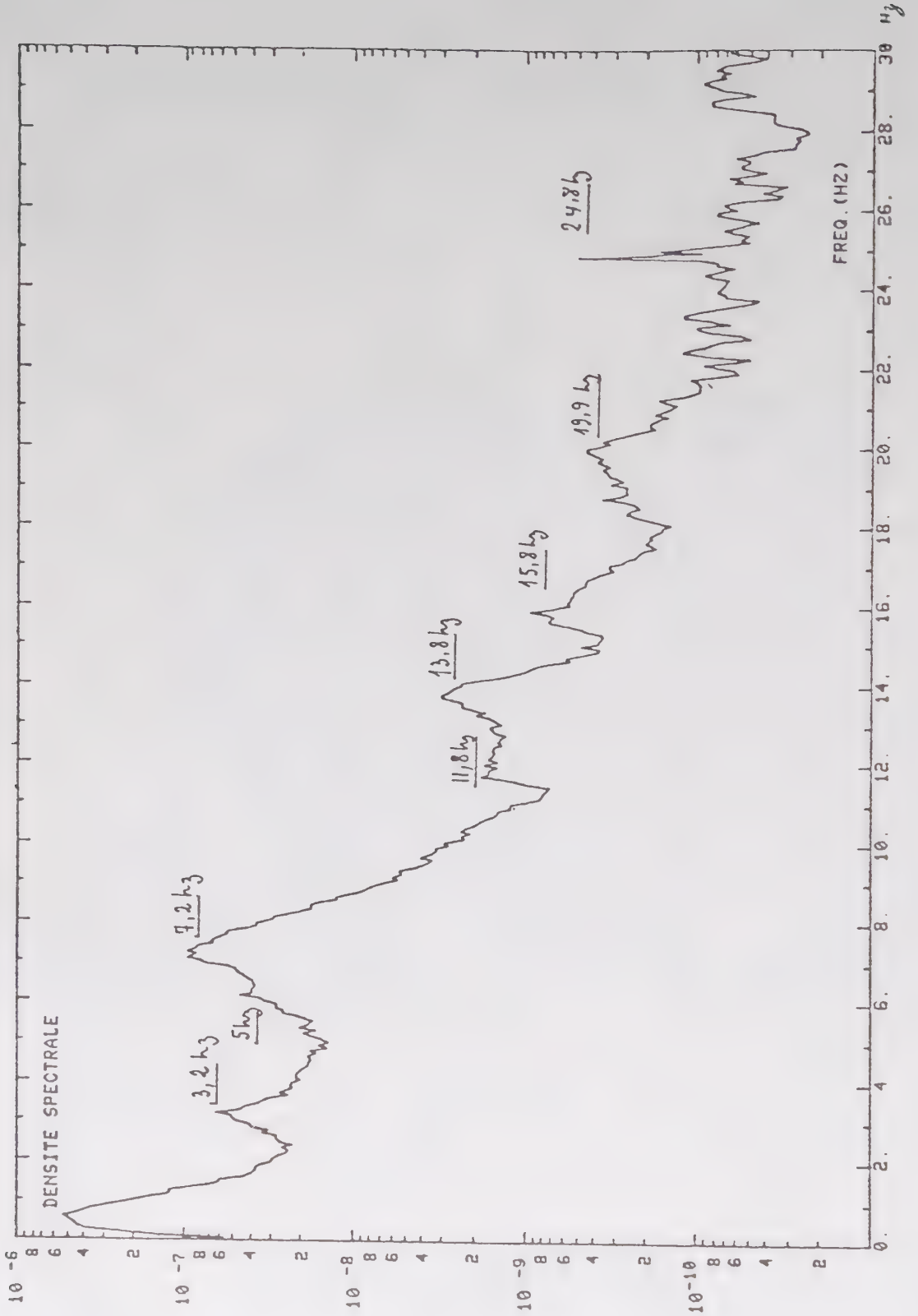


Fig 10: Spectre du bruit neutronique

Fréquences	Chambres neutroniques
3,2 hz	Mouvement des Internes Tendance MODE 1
Vers 6 hz	Mouvement des Internes Tendance MODE 1
7,2 hz	Mouvement principal des Internes MODE 1
11,5 hz	Peut-être MODE 2 des Internes
13,8 hz	Mouvement des Internes Tendance MODE 1
15 - 16 hz	Balancement des Internes dans 2 directions ortho-gonales.
19,5 hz	MODE 2 des Internes

Fig 11: Interprétation des résonnances du spectre

3.1.2. - But de l'étude

Cette étude a été entreprise dans le cadre de la détection précoce d'incidents. Il s'agit de déterminer un algorithme de surveillance pouvant fonctionner en ligne et donner une décision rapidement. La signification d'une décision d'anormalité est à prendre dans le sens : "une tendance anormale dans le comportement vibratoire du coeur vient d'être détectée". Message éventuellement accompagné d'informations concernant cette tendance et destiné à déclencher un processus d'inspection de niveau supérieur. Il s'agit de trouver le maillon de niveau 0 dans un processus de surveillance hiérarchisé schématisé figure 12.

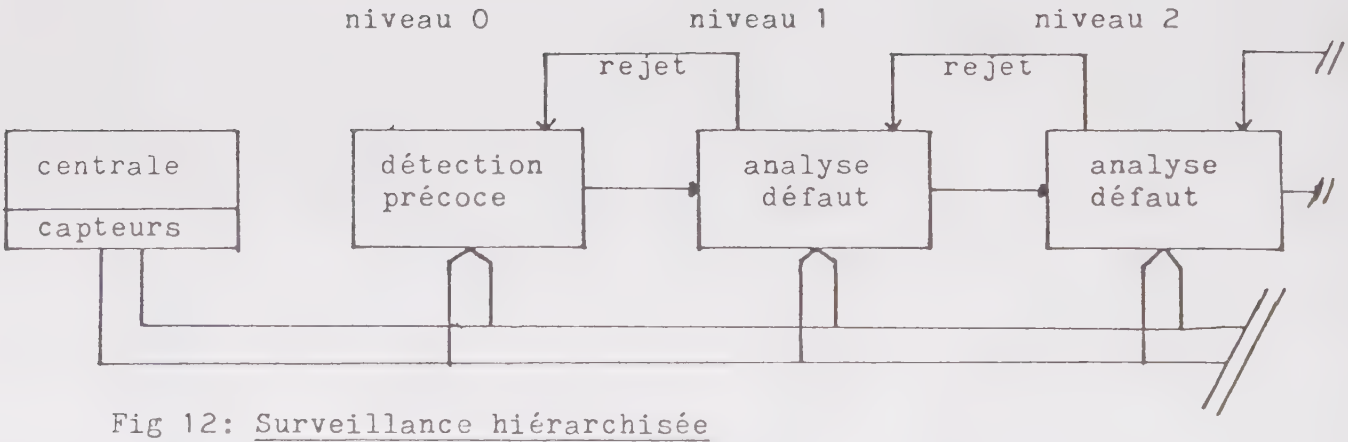


Fig 12: Surveillance hiérarchisée

Un tel processus de surveillance fonctionne de la façon suivante :

L'analyse de niveau n est déclenchée sur une décision d'anormalité de niveau n-1. Elle prend une décision entre transmettre l'alarme au niveau n+1 ou la rejeter. L'intérêt d'une telle structure est de réaliser une allocation dynamique des ressources en fonction des besoins. La complexité, l'importance des moyens mobilisés et le coût de ces différentes analyses augmentent avec le niveau. Les qualités demandées au niveau de la détection précoce sont :

- Une grande fréquence de décision
  A l'échelle d'un réacteur nucléaire cela signifie la prise d'une décision toutes les 10 minutes environ.

- Rapidité de décision
  Chaque décision doit se rapporter à des phénomènes survenus dans les 15 minutes la précédant.

- Faibles moyens requis
  La méthode doit être implantable en ligne et ne nécessiter que du matériel courant.

- Fiabilité des alarmes
  Deux critères servent à déterminer la qualité d'un système de surveillance : la probabilité de fausse alarme, c'est à dire la probabilité pour qu'une observation soit faussement déclarée anormale et la probabilité de non détection, c'est à dire la probabilité pour qu'un fonctionnement anormal ne soit pas reconnu. Pour une méthode de surveillance donnée ces deux critères sont opposés et deux méthodes différentes ne pourront être comparées que pour l'un ou l'autre de ces deux critères fixés. Afin d'assurer la compatibilité du niveau "0" avec les niveaux supérieurs et d'en augmenter la souplesse d'emploi il lui est demandé d'avoir une sensibilité ajustable.

3.1.3. - Principe retenu

Le principe retenu est celui exposé au §.2.4.1., deuxième alinéa. La base d'apprentissage est formée d'enregistrements des signaux, émis par les huit sections de détection pendant des périodes où le réacteur fonctionnait en régime stationnaire. Aucune anormalité n'est apparue pendant ces enregistrements. Nous considérons comme "forme" une durée déterminée d'observation simultanée de ces huit signaux.

Le pré-traitement a pour rôle d'extraire de ces formes les caractéristiques jugées pertinentes relativement au problème de surveillance.

Dans notre cas ces caractéristiques sont les résonances des spectres neutroniques. Chaque forme est alors représentée par le jeu de paramètres extraits définissant ces caractéristiques. L'apprentissage effectue une classification automatique de l'ensemble des formes de la base d'apprentissage en un certain nombre de groupes de telle façon qu'au sein de chaque groupe les formes soient aussi semblables que possible et dissemblables d'un groupe à l'autre. La notion de similitude entre formes est établie à travers des définitions de distances entre vecteurs de paramètres extraits.

Lors de la phase de reconnaissance, une forme inconnue est présentée à l'algorithme qui en extrait un secteur de paramètres de la même façon que lors de l'apprentissage puis l'affecte à l'un des groupes pré-déterminé par l'apprentissage. Cette affectation se fait sur un critère expérimental de vraisemblance qui permet par seuillage la prise de décision d'anormalité.

La méthode de surveillance proposée dans cette étude est une reconnaissance de formes avec apprentissage non supervisé et rejet lors de la phase de reconnaissance.

3.1.4. - Notations

$\mathcal{F}$: Ensemble des formes de la base d'apprentissage.

F : Un jeu de paramètres extraits d'une forme $F = (R_1 \ldots R_{n_F})$ avec n_F le nombre de résonances de F.

R : Un triplet-résonance (fréquence, amplitude, énergie).

f : Une fréquence.

a : Une amplitude (maximum du spectre d'une résonance).

e : Energie d'une résonance.

d, d' : Distances entre triplet-résonances.

D_o, D : Distances entre formes.

ω : classe ou groupe de formes.

T_f : Durée d'une forme.

ρ : Coefficient d'auto-corrélation.

T_c : Fenêtre des fonctions d'auto-corrélation

t_c : Paramètre de la fenêtre en cosinus surélevé.

Fseuil : Seuil utilisé par l'algorithme de réduction.

Pseuil : Seuil utilisé pour déterminer les fréquences de résonance fortement probables.

fseuil : Seuil utilisé par l'algorithme de classification automatique en fréquence.

Zseuil : Seuil utilisé par l'algorithme de séparation.

3.2. - Les données

Les données ayant servi de support à cette étude consistent en des enregistrements analogiques du bruit neutronique issu des deux sections de chacune des quatre chambres du coeur. Le tableau de la figure 13 résume ces données.

Début de cycle : Sept. 80
Fin de cycle : Sept. 81

Date	Durée de l'enregistrement
25.11.80	60 mn
16.12.80	45 mn
15.01.81	45 mn
28.01.81	30 mn
11.02.81	30 mn
11.03.81	30 mn
06.05.81	30 mn

Fig. 13 : Enregistrement du bruit neutronique

Soit au total 4h30 d'enregistrement. Ces mesures ont été effectuées sans amplification préalable et sans filtrage. Une composante continue parasite différente suivant les voies et les acquisitions est introduite par la chaîne de mesure. Toutes ces données sont supposées de type normal. La figure 14 montre environ 3 secondes d'enregistrement sur une voie.

3.3. - Le pré-traitement

Première étape de la chaîne de reconnaissance, le pré-traitement a pour but d'extraire les paramètres jugés pertinents et de réduire leur nombre. Celui qui a été adopté comprend deux parties. Nous avons baptisé la première "pré-traitement à priori"car elle consiste en une transformation pré-déterminée des données. La deuxième est dénommée " à postériori" car elle est soumise à un apprentissage sur la base de données.

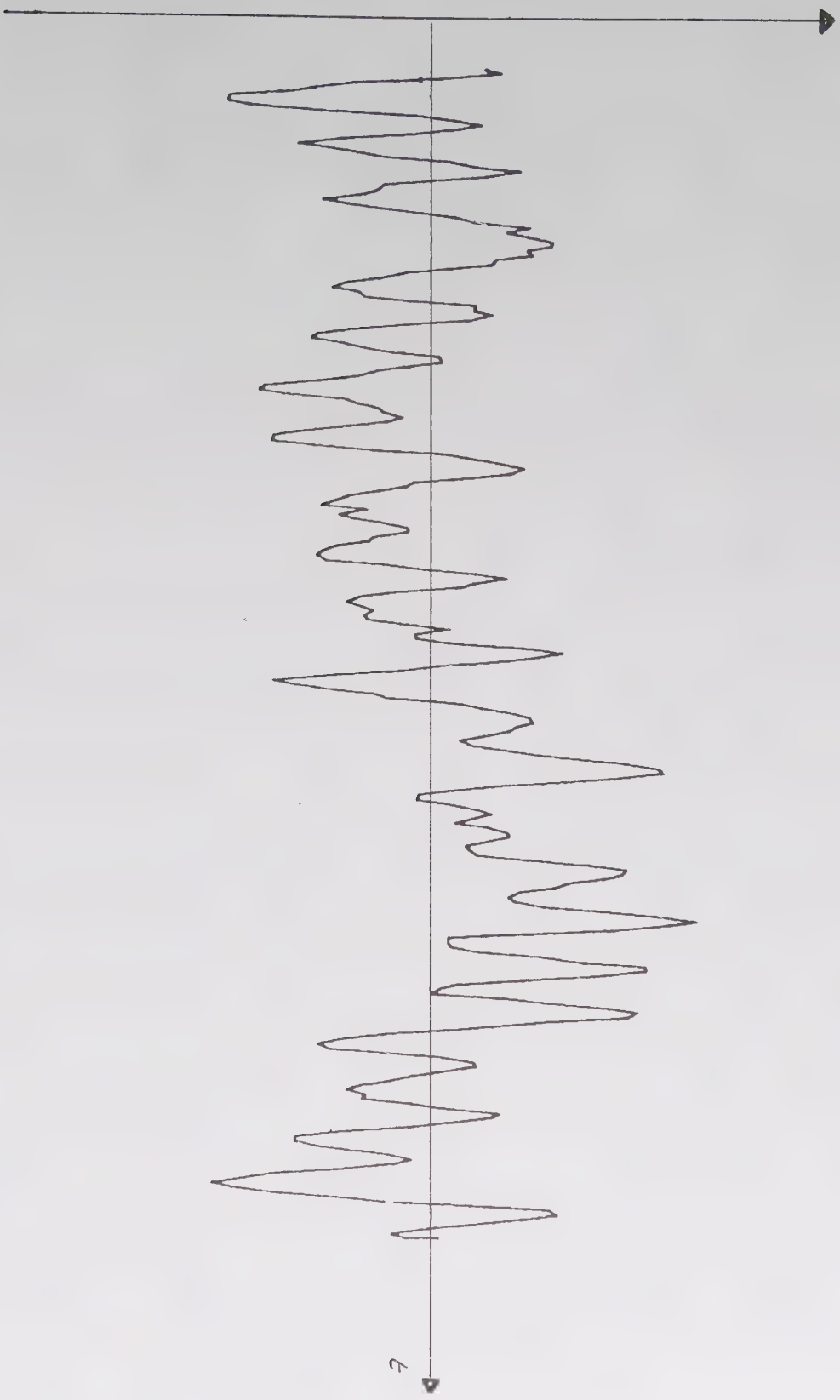


Fig 14: 3s de bruit neutronique

3.3.1. - Pré-traitement à priori

Cette première partie du pré-traitement est schématisée figure 16. Les signaux émis par les 8 sections de détection sont découpés en tranches de durée T_f. Chaque forme est alors représentée par 8 tranches simultanées. Une numérisation est ensuite effectuée à la fréqunece d'échantillonnage de 80 Hz. Le repliement du spectre est négligeable pour notre application ce qui nous dispense de pré-filtrage. Les corrélations sont ensuite estimées de la façon empirique classique :

$$\rho_P = \frac{1}{N} \sum_{i=1}^{N} x_i \, x_{i+p}$$

sur un intervalle temporel T_c. Les corrélations sont pondérées par une fenêtre en cosinus surélevé.

$$F(t) = 1 \qquad t < t_c$$

$$F(t) = (1 + \text{Cos } \pi(t - t_c)/(T_c - t_c))/2 \qquad t > t_c$$

Les spectres sont calculés par transformation de Fourier de 0 à 30 Hz. Les maxima sont détectés ainsi que la surface sous le spectre comprise entre les deux minima encadrant chaque maximum. Une forme est maintenant représentée par huit jeux de triplets représentatifs des maxima des spectres. Chaque triplet étant composé d'une fréquence, d'une amplitude spectrale et d'une énergie. Les quatre chambres subissent le même phénomène physique, il n'est donc pas surprenant de constater une certaine ressemblance entre les spectres extraits. Toutefois, toutes les résonances ne sont pas vues par toutes les sections. D'autre part, un certain nombre de pics détectés ne correspond à aucun phénomène physique et n'est dû qu'à la forme hachée des spectres.

Une classification automatique du type hiérarchisée ascendante avec critère d'agrégation barycentrique est réalisée sur l'ensemble des couples (fréquence, amplitude) affectés de la masse de leur énergie. Les agrégations sont accomplies jusqu'à un écart en fréquence entre centres maximum : Fseuil. La figure 17 illustre le principe de cet algorithme.

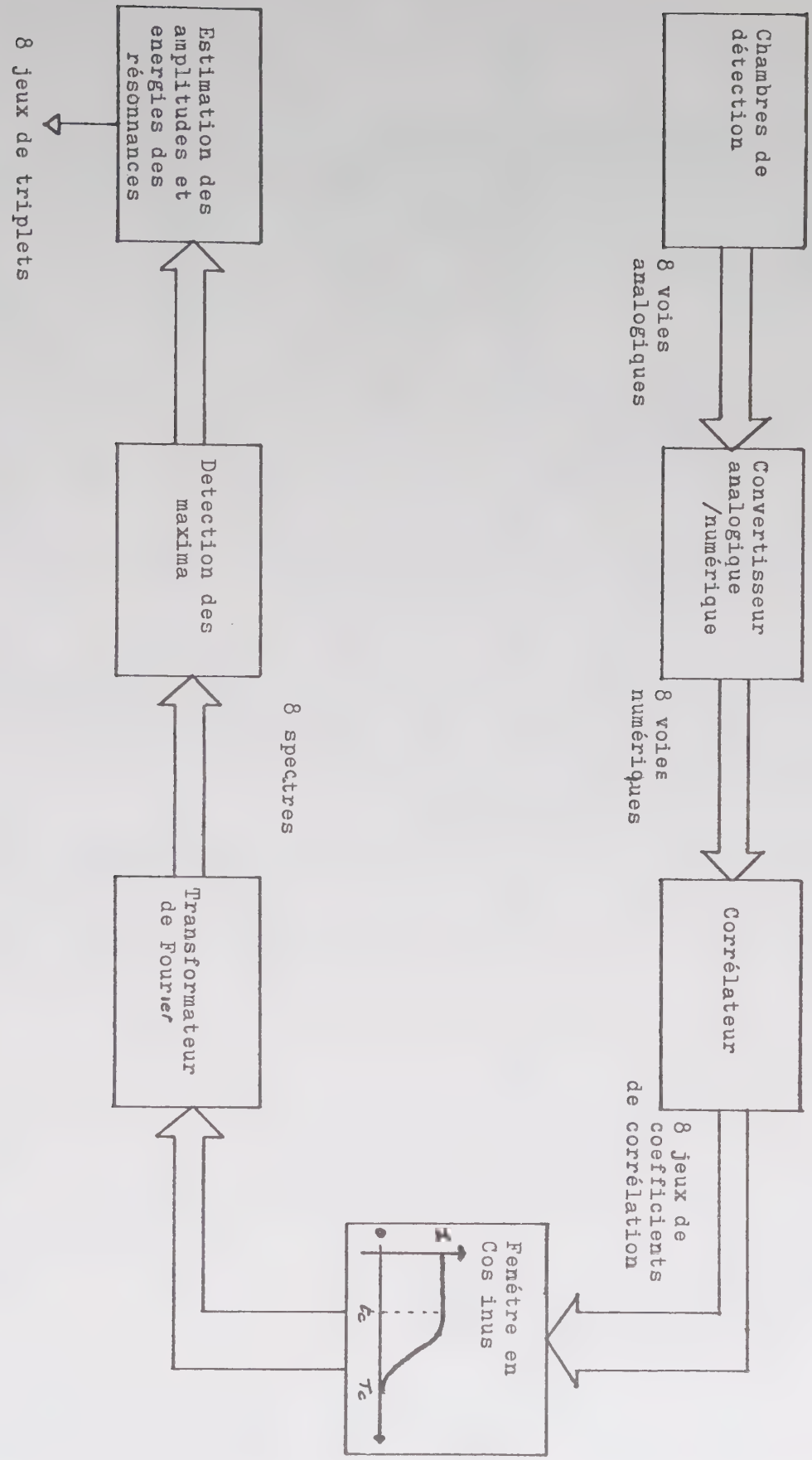
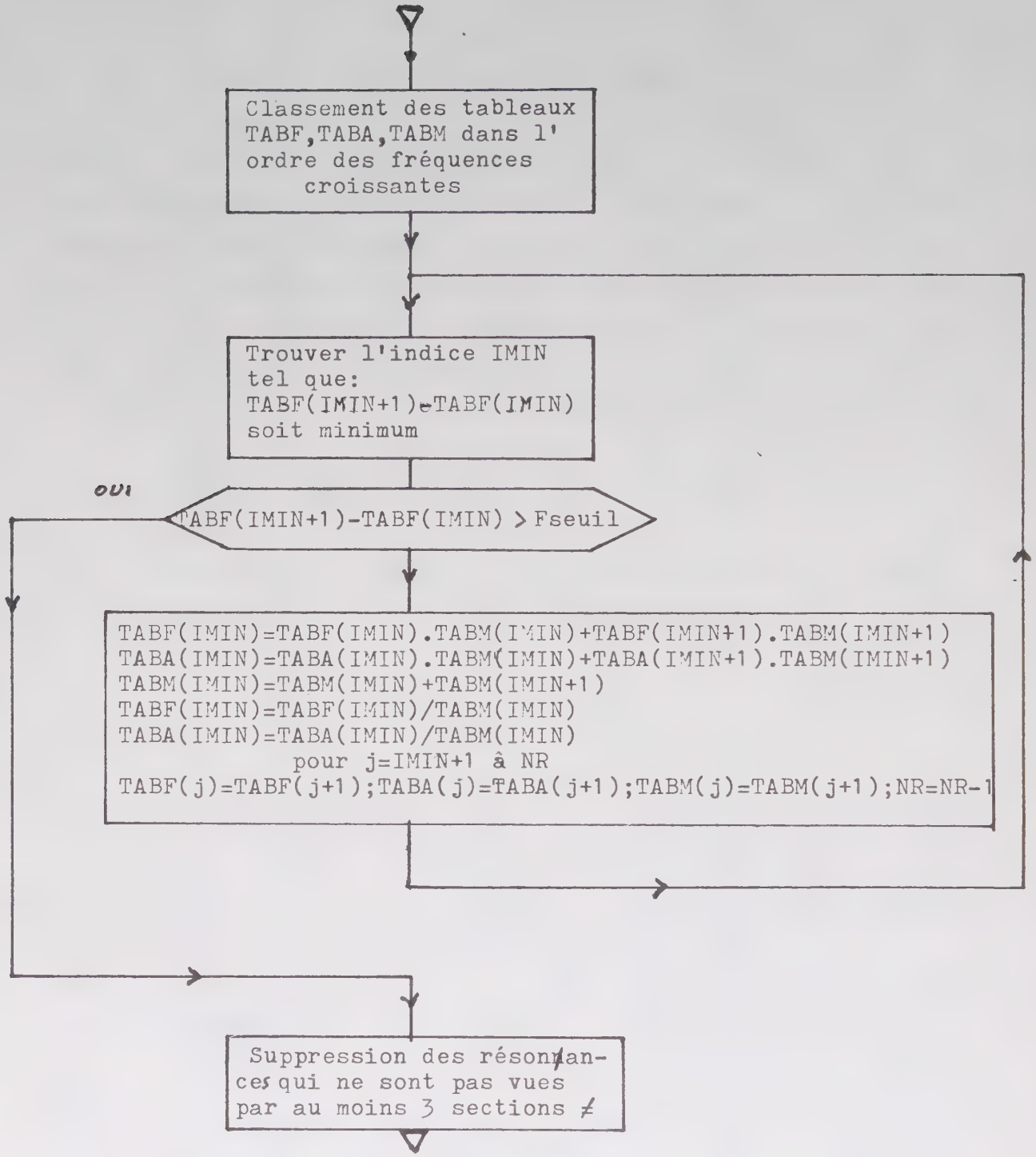


Fig.16 : Schéma du pré-traitement à priori



NR: Nombre de résonnances
TABF: Tableau des fréquences
TABA: Tableau des amplitudes
TABM: Tableau des énergies

Fig 17: Réduction de paramétres par classification automatique

L'ensemble des fréquences, l'ensemble des amplitudes et l'ensemble des énergies correspondant aux maxima détectés sur les huit spectres sont classés dans trois tableaux TABF, TABA et TABM dans l'ordre croissant des fréquences. Les deux fréquences les plus proches sont trouvées. Si l'écart entre elles est inférieur au seuil Fseuil, nous considérons qu'elles proviennent d'une même résonance et remplaçons les caractéristiques des deux maximas correspondant par celles d'un troisième calculées de la façon suivante :

- Fréquence : barycentre des deux fréquences initiales affectées de leur énergie.
- Amplitude : barycentre des deux amplitudes initiales affectées de leur énergie.
- Energie : somme des énergies initiales.

Le processus est reconduit jusqu'à ce que l'écart minimum entre fréquence soit supérieur à Fseuil. Parmi l'ensemble des résonances ainsi déterminées, ne sont conservées que celles dans la détermination desquelles est intervenue l' agrégation d'au moins trois maxima issus de voies différentes.

Chaque forme est maintenant représentée par un seul jeu de triplets dont le nombre dépend d'elle et de la valeur du seuil. Celui-ci doit être choisi suffisamment faible pour interdire des agrégations entre triplets issus de résonances différentes et suffisamment fort pour permettre les agrégations entre triplets représentatifs d'une même résonance. Le tracé de la courbe de variation du nombre moyen de triplets par forme en fonction du seuil peut servir à le déterminer. En effet, lorsque le seuil est inférieur au rayon de la dispersion des triplets autour de la résonance qu'ils représentent cette courbe doit diminuer rapidement pour présenter une partie stationnaire lorsque le seuil devient supérieur à ce "rayon" et rechuter violemment lorsqu'il devient supérieur à l'écart minimum entre résonnances différentes.

Les bornes des valeurs possibles de Fseuil peuvent donc être estimées par les bornes du domaine dans lequel cette courbe est stationnaire. La figure 25 montre une telle courbe obtenue de façon expérimentale.

3.3.2. - Pré-traitement à postériori

L'étape précédente nous fournit une représentation de la forme en un jeu de triplets (position, amplitude, énergie). Elle nécessite la détermination d'un seuil Fseuil. L'écart entre résonances physiques (voir figure 11) n'étant pas constant, il faudrait que ce seuil soit fonction de la fréquence. Plutôt que de chercher à mettre en évidence de façon automatique cette dépendance ou de l'introduire à priori nous avons préféré choisir la plus petite valeur de Fseuil assurant toutefois une certaine réduction. La valeur retenue est celle correspondant au début de la zone stationnaire de la courbe expérimentale de la variation du nombre moyen de (triplets-résonance) par forme en fonction du Fseuil. Le choix de cette valeur minimale nous assure que des associations barycentriques entre (triplet-résonances) représentatifs de résonances physiques différentes ne seront pas accomplies. Exception faite du doublet de fréquences (19, 19.5 Hz) que nous considérons globalement dans cette étude. Par contre tous les "pics" parasites n'ont pas été supprimés. Pour ce faire nous avons utilisé l'algorithme de pré-traitement à postériori suivant :

(1) Estimer la courbe de la variation de la probabilité d'avoir une résonance à la fréquence f en fonction de f par calcul d'un histogramme sur la base d'apprentissage et par lissage.

L'intervalle de fréquence 0-30 Hz est découpé en 300 intervalles élémentaires de 0,1 Hz. L'histogramme est calculé de la façon habituelle : NFO(I) représente le nombre de formes ayant une résonance comprise entre (I-1)/10 et I/10 Hz. Cet histogramme est ensuite lissé et normalisé par la transformation :

$$DFO(I) = \sum_{j=I-NS}^{I+NS} NFO(I)/2NS + 1/NF$$

où NS représente la fenêtre du lissage. (Dans notre application NS = 2).

et NF le nombre de formes de la base d'apprentissage.

La figure 18.1. montre une courbe ainsi obtenue.

(2) Séparation du domaine fréquentiel en deux sous-domaines correspondant à l'ensemble des fréquences considérées comme fortement probables et à l'ensemble des fréquences considérées comme faiblement probables.

(1) - Trouver la plus grande valeur de DFO(I) pour I variant entre 1 et 300 supérieure au seuil Pseuil et différente de 1. Soit Imax l'indice correspondant. Si aucune valeur de DFO(I) ne satisfait ces conditions aller en (4).

(2) - Trouver la plus petite valeur de I supérieure à Imax telle que DFO(I) > DFO(I-1) soit I2 cette valeur.

- Trouver la plus grande valeur de I inférieure à Imax telle que DFO(I) > DFO(I+1) soit I1 cette valeur.

(3) - Pour tous les indices compris entre ces deux valeurs incluses poser DFO(I) = 1. Aller en (1).

(4) Pour chaque indice I si DFO(I) est différent de 1 poser DFO(I) = 0.

L'ensemble des fréquences considérées comme fortement probable est défini comme l'ensemble des fréquences comprises dans un intervalle élementaire I tel que DFO(I) =1.

La figure 18.2. montre le partitionnement obtenu à partir de la figure 18.1 et pour une valeur de Pseuil de 0,3. Les zones repérées sur l'axe p = 0 correspondent à des fréquences considérées comme faiblement probables, celles repérées sur l'axe p = 1 aux autres.

(3) Mémoriser les deux sous ensembles déterminés par l'étape (2) afin de pouvoir effectuer la transformation décrite à l'étape 4 lors de la phase de reconnaissance.

Cette mémorisation est effectuée à travers celle du vecteur DFO résultant de l'étape (2).

(4) Transformer chaque forme de la base d'apprentissage de la façon décrite ci-dessous.

Soit F la forme à transformer

$F = (R_1, .. R_{n_F}) = ((f_1\ a_1, e_1), ...(f_{n_F}, a_{n_f}, e_{n_F}))$

On suppose $0 < f_1 < f_2 .. < f_{n_F}$

(1) Poser j = 0 j = indice de la résonance en cours de traitement.

(2) Incrémenter j par 1
Si j est supérieur à n_F-1 aller en (4)

(3) Considérer le couple de résonances R_j, R_{j+1}
Si $f_{j+1} - f_j > 2$ x Fseuil aller en (2)
Si f_j et f_{j+1} sont fortement probables aller en (2)
Calculer $f = (f_j\ e_j + f_{J+1}\ e_{j+1})/(e_j + e_{j+1})$
Si f est faiblement probable aller en (2)
sinon + poser - $f_j = f$

- $a_j = (a_j\ e_j + a_{j+1}\ e_{j+1})/(e_j + e_{j+1})$

- $e_j = e_j + e_{j+1}$

+ supprimer le triplet R_{j+1} et réordonner la forme
+ faire $n_F = n_{F-1}$, $j = \max(j-2, 0)$ et aller en (2).

(4) FIN.

⑤ identique au ① mais les calculs sont effectués sur la base transformées par l'étape ④ la figure 18.3 montre la courbe obtenue

⑥ - ⑦ - ⑧ - identiques aux ② - ③ - ④ -

Le critère que nous avons retenu pour juger de l'efficacité de cette méthode de réduction avec apprentissage est l'entropie de la courbe de la variation de la probabilité qu'une forme admette une résonance à la fréquence f en fonction de f :

$$H = \int - P(f) \log P(f) \, df.$$

Il apparait que cette entropie décroit fortement lors du premier passage (étapes ① à ④), faiblement lors du second (étapes ⑤ à ⑧) et devient quasi-stationnaire lors de passages ultérieurs. Nous nous sommes donc limités à deux passages. Cet algorithme de réduction à postériori est la continuation de celui décrit au paragraphe 3.3.2. La décision de réaliser un regroupement barycentrique n'y est plus prise sur un critère d'écart en fréquence mais sur celui de supression de résonances considérées comme faiblement probables. Le choix du paramètre Pseuil n'est pas critique il apparait clairement sur la figure 18.2. que toute valeur de Pseuil comprise entre 0,2 et 0,5 eût donné les mêmes résultats. L'étape suivante ne fait pas partie du prétraitement. Ses résultats seront utiles aux algorithmes de classification automatique utilisés par la suite.

⑨ La courbe de la variation de la probabilité d'avoir une résonance f en fonction de f est à nouveau calculée (sur la base d'apprentissage ayant été transformée deux fois). Le domaine fréquentiel est séparé en bandes dont les bornes sont les milieux de chaque sous domaine connexe considéré comme faiblement probable. La figure 18.4 montre les bandes obtenues sur l'exemple traité. Remarquons que cette segmentation n'est pas faite résonance par résonance. Les zones considérées comme faiblement probables sont repérées par un trait large sur l'axe P = 0, les bornes des bandes fréquentielles par les traits verticaux.

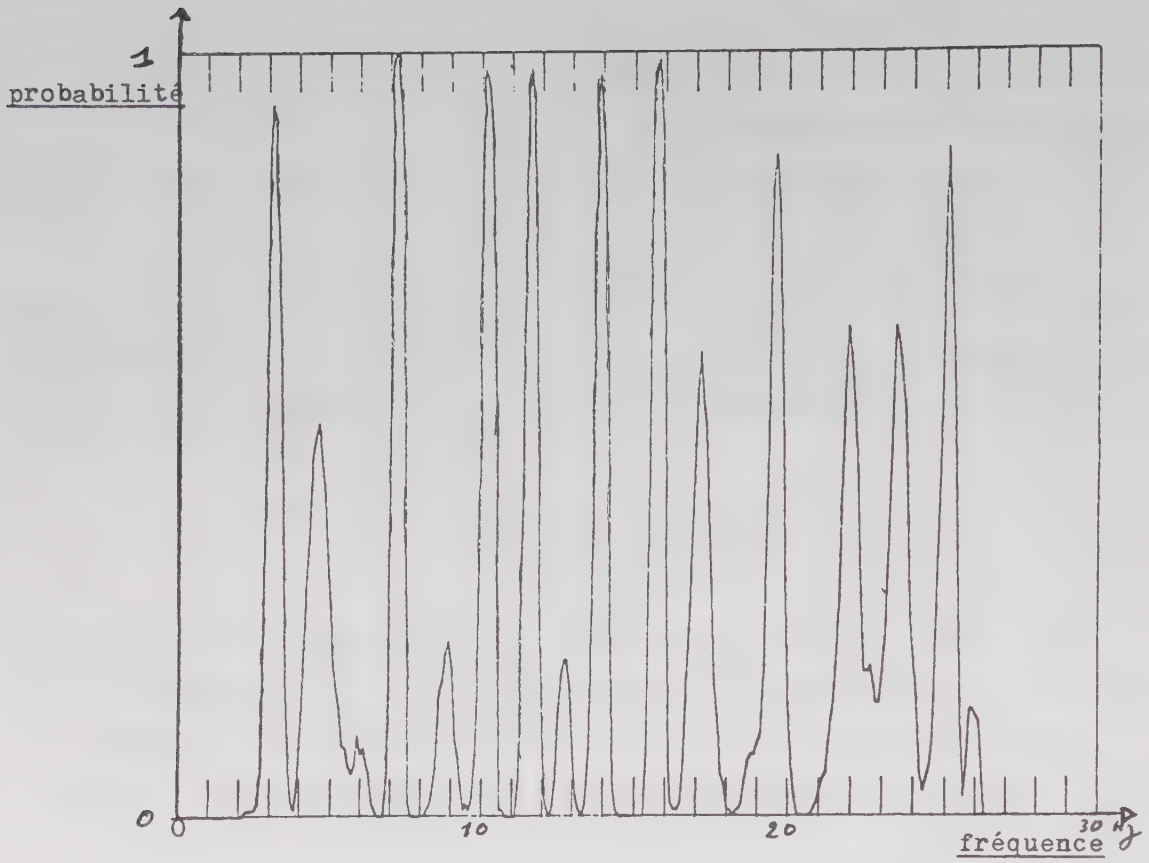


FIG 18-1: Probabilité qu'une forme admette une résonance à la fréquence f en fonction de f

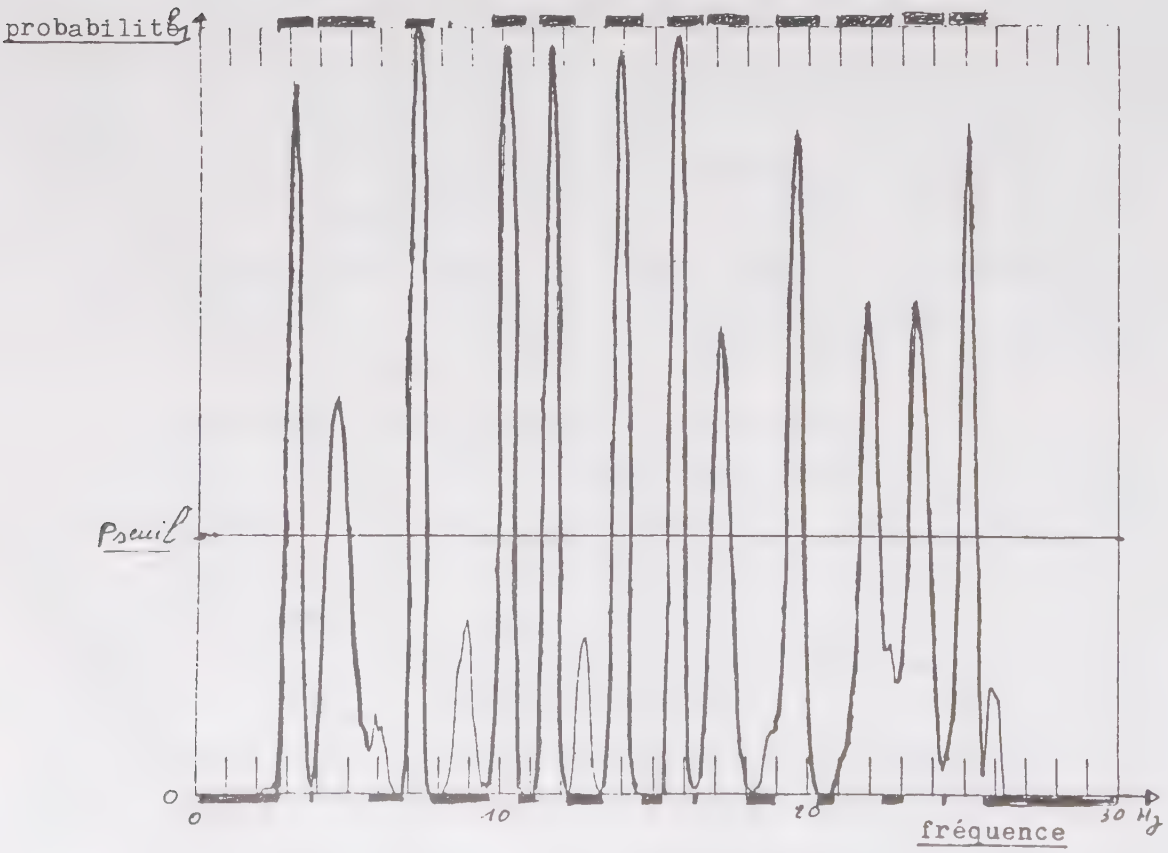


FIG 18-2: Classification fortement/faiblement probable

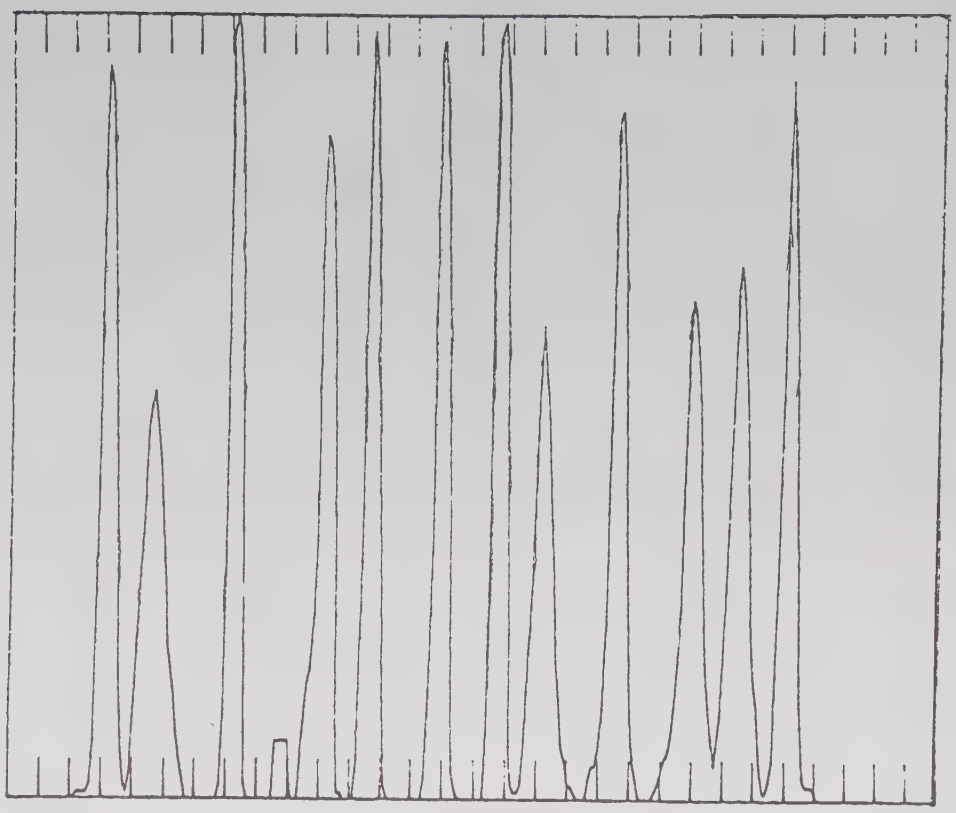


Fig 18-3: Probabilité qu'une forme de la base d'apprentissage transformée une fois admette une résonnance à la fréquence f

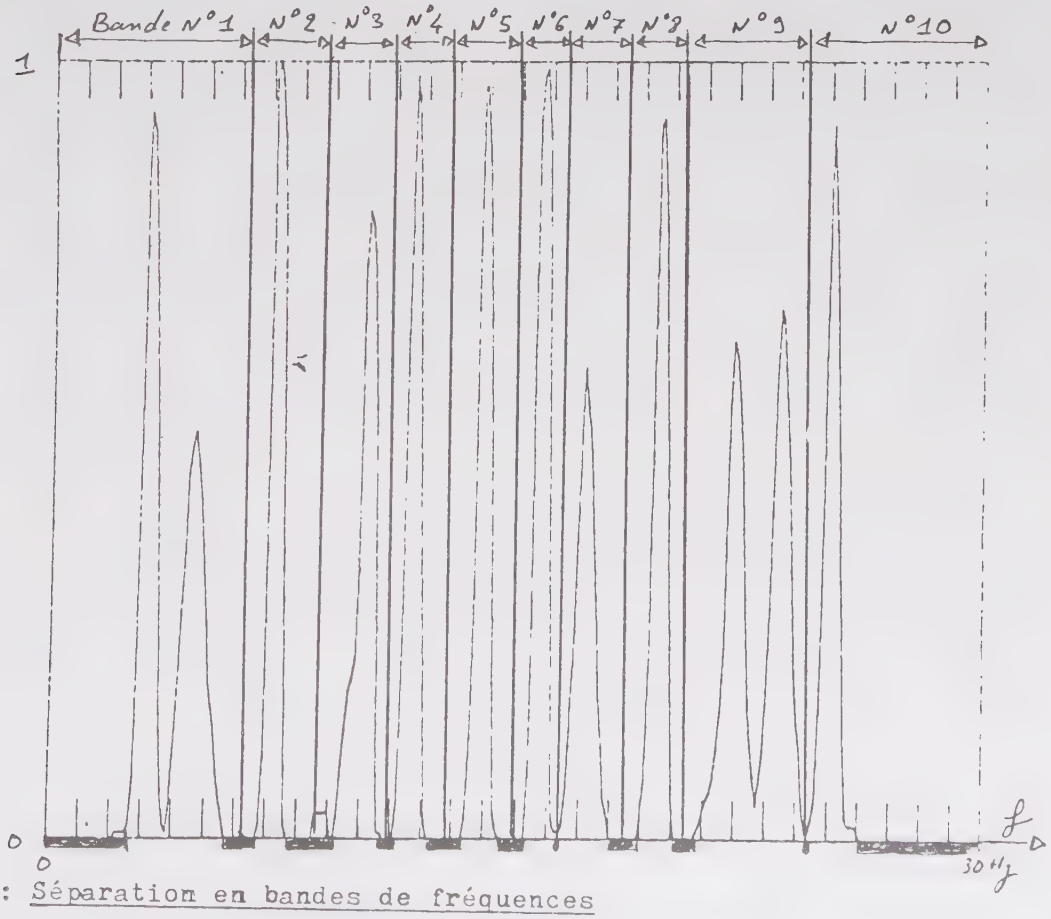


Fig 18-4: Séparation en bandes de fréquences

3.3.3. - Commentaire sur la réduction

Pour assurer la faisabilité du traitement nous sommes amenés à réduire le nombre des paramètres extraits en concentrant sur eux le maximum d'informations. Mais sur quels critères doit s'effectuer cette réduction ? Faut'il garder les paramètres variant faiblement parmi les formes d'apprentissage ou au contraire ceux admettant une grande dispersion ? Dans le cas d'un problème de reconnaissance classique où la base d'apprentissage contient des épreuves de toutes les classes que l'on veut identifier ,le choix est en général possible (analyse en composantes principales, analyse factorielle, critères représentatifs et/ou discriminants). Dans le cas présent où l'apprentissage ne travaille que sur des formes d'un seul type (normal) il ne peut y avoir de réponse stricte : les paramètres du premier groupe admettant une faible dispersion parmi les formes normales sont ceux qui sont représentatifs de cette normalité. Par contre, ils ne portent pas d'informations relativement aux différents types de normalité que nous cherchons à mettre en évidence. Ces informations étant portées par les paramètres du second groupe. La méthode de réduction que nous avons utilisée est basée sur des regroupements barycentriques successifs entre (triplets-résonance) voisins dans une même forme. Cette méthode de regroupement présente l'avantage de conserver globalement l'énergie. Celle-ci est simplement répartie de façon différente entre les résonances avant et après traitement. D'autre part, l'"énergie" de chaque triplet est estimée par l'intégrale du spectre entre les abscisses des minima encadrant chaque maximum. Elle mesure donc bien le crédit à apporter au fait d'avoir pris en compte ce maximum. La règle d'association barycentrique entre maxima pondérés par leur énergie est donc intuitivement satisfaisante.

Les fréquences de résonance mises en évidence par le pré-traitement ont les valeurs moyennes (elles varient légèrement d'une forme à l'autre) suivantes : 3,2 Hz/4,6 Hz/ 7,2 Hz/10 Hz/11,5 Hz/13,8 Hz/15,7 Hz/17 Hz/19,5 Hz/21,9 Hz/ 23,4 Hz/24,9 Hz/.

Nous retrouvons les fréquences mises en évidence dans les rapports [29-30] et rappelées figure 11 ainsi qu'un certain nombre de fréquences supplémentaires.

Il est intéressant de constater (fig. 18.4.) que les fréquences figurant au tableau de la figure 11 sont considérées par notre algorithme comme les plus probables. Le reproche souvent émis à l'encontre des algorithmes de nature heuristique est l'utilisation de seuils arbitraires. Les seuils utilisés dans cette étude sont soit déterminés automatiquement soit expérimentalement par l'étude de variations de courbes.

3.4. - Classification automatique

Le but poursuivi par cette étape est le partitionnement de la base d'apprentissage en classes de telle façon qu'au sein de chaque classe les formes extraites soient aussi semblables que possible et dissemblables d'une classe à l'autre. La classification retenue comprend trois étapes. La première est une classification automatique en fréquence. La similitude entre formes y est mesurée par les positions respectives des résonances les composant. La deuxième étape considère chacune des classes précédemment trouvée et détermine si elle doit être divisée ou non sur un critère de dispersion en distance. Cette étape réalise l'apprentissage d'une distance inter-forme. La troisième étape est l'algorithme des K moyennes initialisé par la classification obtenue où la distance euclidienne a été remplacée par la distance déterminée ci-dessus.

3.4.1. - Classification automatique en fréquence

La méthode utilisée est une variante de l'algorithme du rayon maximal. Son principe est exposé ci-dessous.

(1) Initialiser le nombre de classes à 0.

(2) Trouver une forme non affectée, non marquée. S'il n'existe plus de telles formes aller en 6 . Incrémenter le nombre de classes par 1. Initialiser le centre de la nouvelle classe par cette forme.

(3) Considérer successivement toutes les formes non affectées. Si la forme a même nombre de triplets que le centre du groupe en cours de détermination et que la plus grande différence entre position des résonances de même ordre est inférieure à fseuil, affecter la forme au groupe et calculer le nouveau centre.

(4) Si il y a eu une affectation lors de l'étape 3 y retourner.

(5) Si le groupe ne possède qu'une forme, marquer cette forme, décrémenter le nombre de groupes et aller en (2) sinon aller directement en (2).

(6) Considérer les formes non affectées. Affecter ces formes sur le critère de la distance D_o aux centres déterminés minimale.

Le nombre de groupes ainsi générés dépend du choix du paramètre fseuil. Ce choix n'est pas arbitraire mais est effectué de la façon automatique suivante ; si les groupes à une forme étaient autorisés, la courbe écart moyen en fréquence d'une forme au centre auquel elle est affectée serait strictement croissante en fonction de fseuil. Le fait d'interdire de tels groupes et d'affecter les formes concernées sur un critère de distance faisant intervenir les amplitudes et les énergies donne à cette courbe un minimum. Le paramètre fseuil est ajusté automatiquement pour réaliser ce minimum.

La distance D_o est définie au §.3.4.2.

A la fin de cette classification automatique sur critère d'écart en fréquence, les formes de la base d'apprentissage sont réparties en un certain nombre de groupes. Quelques formes peuvent rester isolées. Le nombre de ces formes est peu important et décroit avec la taille de la base d'apprentissage. La courbe expérimentale de la variation du nombre de formes non affectée à un groupe en fonction du nombre de formes de la base d'apprentissage est donnée figure 29.

Lors d'un apprentissage sur 240 formes seules deux d'entre elles ne furent affectées à aucun groupe. Il est apparut après analyse que les signaux ayant servi au calcul des paramètres de ces formes présentaient un affaiblissement anormal.

Nous considérons ces formes comme marginales et les retirons de la base d'apprentissage. Les classes déterminées par cette première partie de l'apprentissage possédent les propriétés suivantes :

- Toutes les formes d'une même classe ont même nombre de triplet-résonances,

- L'écart en fréquence entre résonances de même ordre des formes d'une même classe est faible (de l'ordre de fseuil).

Afin d'illustrer le fonctionnement de cet algorithme nous présentons ci-dessous un exemple sur des données artificiellles. Supposons que la base d'apprentissage soit $B = \{F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8\}$.

F_1 = ((1,3,10), (4,2,5), (6,2,5), (8,3,9))

F_2 = ((1,2,4,11), (5,3,7), (6,3,8), (8,3,9))

F_3 = ((1,1,3,9), (3,8,3,7), (6,3,2,4), (8,1,4,10))

F_4 = ((1,4,10), (5,2,5,8),(6,1,5,9), (7,8,4,10))

F_5 = ((1,4,10), (5,2,5,8),(7,5,4,11))

F_6 = ((1,2,4,11), (5,3,7), (8,3,9))

F_7 = ((1,3,10), (4,2,5), (8,3,9))

F_8 = (1,1,3,9), (3,7,3,7)

et supposons que fseuil = 0,6.

Etape (2) : $\omega_1 = \{F_1\}$ centre C_1 = ((1,3,10),(4,2,5),(6,2,5)
(8,3,9))

Etape (3) : F_2 a même nombre de triplets que le centre de ω_1 mais l'écart en fréquence sur le 2ème triplet est supérieur à fseuil donc F_2 n'est pas affectée à ω_1
F_3 a même nombre de triplets que ω_1 et l'écart maximum en fréquence est inférieur à fseuil,
F_3 est affectée àla classe ω_1 $\omega_1 = \{F_1, F_3\}$
le centre de ω_1 est recalculé.
$C_1 = (F_1+F_3)/2 = ((1,05,3,9,5),(3,9,2,5,6),(6,15,2,4,5),(8,05,3,05,9,5))$
F_4, F_5, F_6, F_7 et F_8 ne peuvent être affectés à ω_1.

Etape (4) : Il y a une affectation lors de l'étape 3 celle-ci est donc à nouveau relancée.

Etape (3) : Aucune nouvelle affectation n'est possible,

Etape (4) : On passe au (5).

Etape (5) : Le groupe posséde deux formes, il est donc conservé.

Etape (2) : $\omega_2 = \{F_2\}$ $C_2 = ((1,2,4,11),(5,3,7),(6,3,8),(8,3,9))$

Etape (3) : F_1, F_2, F_3 sont déjà affectées
F_4 est affectée à ω_2
donc $\omega_2 = \{F_2, F_4\}$
$C_2 = ((1.1,4,10.5),(5.1,4,7.5),(6.05,4,8.5),(7.9,3.5,9.5))$
F_5 F_6, F_7 F_8 n'ont pas même nombre de triplets que C_2.

Etape (4) : Il y a eu une affectation de l'étape (3) celle-ci, est donc relancée.

Etape (3) : $F_1 \rightarrow F_4$ sont déjà affectées
$F_5 \rightarrow F_8$ sont inaffectables à C_2.

Etape (4) : On passe en (5)

Etape (5) : Le groupe posséde deux formes il est donc conservé.

Etape (2): $\omega_3 = \{F_5\}$ $C_3 = ((1,4,10),(5.2\ 5,8),(7.5,4,11))$

Etape (3): F_1 à F_5 sont déjà affectées
F_6 est affectable
$\omega_3 = \{F_5, F_6\}$ $C_3 = ((1.1,4,10.5),(5.1,4,7.5),$
$(7.75,3.5,10))$
F_7 a même nombre de triplets que C_3 mais l'écart en fréquence sur le 2ème triplet est supérieur à fseuil.
F_8 n'a pas même nombre de triplets que C_3.

Etape (4): Retour étape 3.

Etape (3): Pas de nouvelle affectation.

Etape (4): Passage étape 5.

Etape (5): Le groupe posséde deux formes, il est donc conservé.

Etape (2): $\omega_4 = \{F_7\}$ $C_4 = ((1,3.,10),(4.,2.,5.),(8.,3.,9.))$

Etape (3): F_1 à F_7 sont déjà affectées
F_8 n'a pas même nombre de triplets que C_4.

Etape (4): Passage étape 5.

Etape (5): Le groupe ne possède qu'une forme, il est donc supprimé, la forme F_7 est marquée.

Etape (2): La première forme disponible est F_7 mais cette forme est marquée donc $\omega_4 = \{F_8\}$
$C_4 = ((1.1,3.,9.),(3.7,3.,7.))$

Etape (3): Les formes F_1 à F_6 sont déjà affectées.
La forme F_7 n'a pas même nombre de triplets que C_4.

Etape (4): Passage étape 5.

Etape (5): Le groupe ne posséde qu'une forme, il est donc supprimé. La forme F_8 est marquée

Etape (2): Il n'existe plus de formes non marquées et non affectées donc passage à l'étape (6).

Etape (6): Supposons que la distance D_o soit définie par
$D_o(F_i, F_j) = \infty$ si F_i et F_j n'ont pas le même nombre de résonances.
$D_o(F_i, F_j)$ = Distance Euclidienne entre F_i et F_j sinon, F_7 serait affectée au groupe ω_3 dont le centre serait recalculé.
F_8 ne serait pas affectable. F_8 est donc retirée de la base d'apprentissage.

3.4.2. - Distances inter-résonance, distances inter-formes

- Distances inter-résonance

Chaque résonance est représentée par un triplet (fréquence, amplitude maximale du spectre, énergie) (f, a, e). En supposant chaque résonance issue d'un oscillateur harmonique amorti son spectre peut être approximé par :

$$R(x) = \frac{uaf^2}{(x^2-f^2)^2 + ux^2} \quad \text{avec } u = \alpha \left(\frac{e}{a}\right)^2$$

et α un paramètre d'approximation que nous avons fixé à 4/3.

Deux distances inter-résonance ont été utilisées.

$$d(R_1,R_2) = \int |R_1-R_2|(x)\,dx \Big/ \int (R_1+R_2)(x)\,dx$$

et

$$d'(R_1,R_2) = \int |\log R_1 - \log R_2|(x)\,dx \Big/ \left[\int |\log R_1| + |\log R_2|(x)\,dx - A\right]$$

où A est une constante.

- Distance inter-forme

Une fois choisie une distance inter-résonance d on définit une distance inter-forme de la façon suivante :

Soient deux formes F_1 et F_2

$F_1 = (R_1^1 - - R_{P_1}^1)$

$F_2 = (R_1^2 \quad R_{P_2}^2)$

si $P_1 \neq P_2$ $\quad D(F_1, F_2) = \infty$

si $P_1 = P_2 = P$ $\quad D(F_1, F_2) = \sum_{i=1}^{P} d(R_i^1, R_i^2)/A_k$

avec A_k des coefficients de pondération. Au § 3.4.1. ces coefficients sont supposés égaux, par la suite A_k est une estimation de l'écart moyen de la distance inter-résonance dans la bande de fréquence (cf. §.3.3.2.) d'appartenance des résonaces R_i^1 et R_i^2.

3.4.3. - Classification automatique sur critère de distance

L'étape précédente nous fourni une séparation de la base d'apprentissage en classes de telle façon que les formes d'une même classe aient même nombre et à peu près mêmes positions de résonances. Par contre aucun critère n'a été retenu quant à la dispersion en amplitude et en énergie de chaque résonance. L'algorithme décrit ci-dessous reconsidère chaque classe déterminée comme indiqué au paragraphe 3.4.1. et analyse résonance par résonance les dispersions de la distance inter-résonance choisie. Les formes d'une même classe ayant des fréquences voisines cette distance mesure en fait des écarts en amplitude ou en énergie.

① - Calcul pour chaque classe des distances moyennes des triplets-résonances des formes de la classe aux triplets-résonances du centre de cette classe

$$E_j^i = \frac{1}{N_i} \sum_{k=1}^{N_i} d(R_j^k, R_j^i)$$

E_j^i = Distance inter-résonance moyenne entre la résonance n° j d'une forme de la classe ω_i et la résonance n° j du centre de la classe ω_i.

N_i = Nombre de formes de la classe ω_i.

R_j^k = jième résonance de la kième forme de la classe ω_i.

$R_j^i \hat{=} \sum_{k=1}^{N_i} R_j^k$ = jième résonance du centre de la classe ω_i.

(2) - Calcul pour chaque bande fréquentielle (définies au paragraphe 3.3.2.) de la distance moyenne des triplets-résonances appartenant à cette bande aux triplets-résonances de même indice du centre de la classe d'appartenance des formes les contenant.

$$A_k = \frac{1}{N_k} \sum_{\substack{\text{formes F} \\ \text{possédant un triplet résonance } R_j^F \\ \text{dans la bande fréquentielle n° k}}} d\,(R_j^F, R_j^C)$$

A_k = Distance inter-résonance moyenne dans la $k^{ième}$ bande fréquentielle

N_k = Nombre de formes possédant un triplet-résonance dans la $k^{ième}$ bande fréquentielle

R_j^R = Triplet-résonance de la forme F compris dans la bande.

R_j^C = Triplet-résonance de même indice que R_j^F du centre de la classe d'appartenance de F.

Si l'on suppose que les triplets-résonance de même indice des formes d'une même classe appartiennent tous à une même bande fréquentielle. (Cette supposition est expérimentalement toujours vérifiée), Les coefficients A_k peuvent être calculés à partir des coefficients E_j^i par la formule suivante :

$$A_k = \frac{1}{N_k} \sum N_i \quad E_j^i$$

l'ensemble des classes ω^i dont le centre posséde un triplet résonance (j) dans la $k^{ième}$ bande fréquentielle.

N_i = Nombre de formes ω_i

N_k = Nombre total des formes concernées.

(3) - Considérer chaque classe ω_i. Si pour l'une des résonances de la classe, la valeur moyenne de la distance inter-résonance calculée pour cette résonance et pour cette classe est supérieure au produit de Zseuil par la valeur moyenne de la distance inter-résonance calculée dans la bande fréquentielle d'appartenance de cette résonance. Séparer cette classe en deux par la méthode décrite ci-après. La conditions de séparation s'écrit simplement :

$E_j^i > \text{Zseuil} \times A_{k(R_j^i)} \longrightarrow$ séparation

avec $k(R_j^i)$ l'indice de la bande de fréquence d'appartenance de R_j^i.

L'algorithme de séparation que nous avons utilisé est une variante de la méthode des k moyennes qui sera décrite au paragraphe 3.4.4.

(1) Trouver les deux formes de la classe ω_i les plus éloignées suivant la résonance j.

F^1 et F^2 $\quad F^1 = (R_1^1 \text{ -- } R_n^1)$

$F^2 = (R_1^2 \text{ -- } R_n^2)$

$d(R_j^1, R_j^2) \geq d(R_j^k, R_j^\ell) \quad \forall F^\ell F^k \in \omega_i$

Initialiser les centres C^1 et C^2 des deux sous classes S^1 et S^2 à générer par ces deux formes.

$C^1 = (R_1^1 \text{ -- } R_j^1)$

$C^2 = (R_1^2 \text{ -- } R_j^2)$

(2) Répartir les formes de la classe ω_i parmi les deux sous-classes avec la règle suivante :

$d(R_j^\ell, R_j^1) \leq d(R_j^\ell, R_j^2) \longrightarrow F^\ell \in S^1$

$d(R_j^\ell, R_j^2) < d(R_j^\ell, R_j^1) \longrightarrow F^\ell \in S^2$

(3) Recalculer les centres des deux sous-classes

$$C^1 = \frac{1}{\text{Nombre de formes de } S^1} \sum_{F \in S^1} F$$

$$C^2 = \frac{1}{\text{Nombre de formes de } S^2} \sum_{F \in S^2} F$$

(4) Si l'étape (3) a modifié la position des centres des sous classes retourner en (2).

(5) Supprimer la classe ω_i et la remplacer par les deux classes S^1 et S^2.

(4) - Si une classe a été séparée lors de l'étape 3 retourner en 1.

(5) - FIN.

3.4.4. - Algorithme des k moyennes

L'algorithme des k moyennes présenté par [BALL] permet lorsque le nombre de classes est connu de déterminer la partition optimale du point de vue de l'écart quadratique moyen d'une forme au centre du groupe auquel elle appartient. (Chaque forme doit être un vecteur de $\mathbb{R}^n$ et la distance utilisée est la norme Euclidienne). Cet algorithme entre dans le cadre plus général des algorihmes des nuées dynamiques [DIDAY] et fonctionne suivant le schéma itératif de la figure 19.

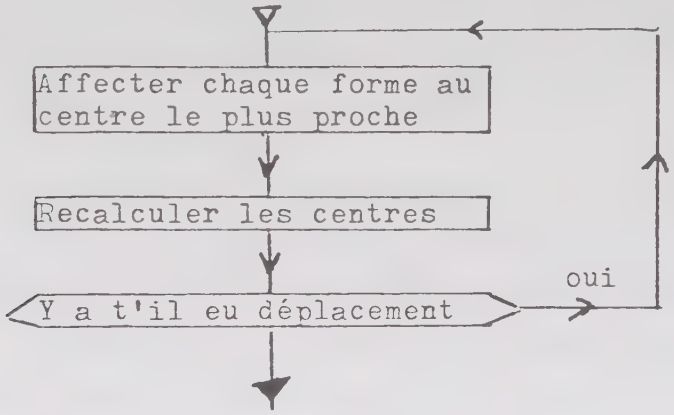


Fig 19: Algoritme des K moyennes

La dernière étape de l'apprentissage est semblable à cet algorithme dans lequel la distance Euclidienne est remplacée par la distance inter-forme D définie au paragraphe 3.4.2. dont les coefficients de pondération ont été estimés lors de la classification automatique sur critère de distance. L'algorithme est initialisé par la donnée des centres des classes déterminées au paragraphe 3.4.3. Notons que le centre d'une classe ne minimise pas pour une répartition donnée l'écart quadratique moyen au sens de notre distance. La convergence vers un minimum relatif de ce critère n'est donc pas assurée. Toutefois, les distances inter-formes proposées sont des fonctions croissantes des écarts en fréquence, amplitude et énergies sur chacune des résonances. La minimisation des écarts quadratiques moyens sur ces caractéristiques doit donc permettre celle des écarts en distance.

Soit C^1 .. C^N les centres des classes trouvées après les méthodes de classification exposées aux paragraphes 3.4.1. et 3.4.3.

(1) Affecter chaque forme de la base d'apprentissage à la classe suivant la règle.

$$D(F, C^i) \leq D(F, C^j) \quad \forall j \neq i \longrightarrow F \in \omega_i$$

(2) Recalculer les centres des classes.

$$\forall i \quad C^i = \frac{1}{Card(\omega_i)} \sum_{F \in \omega_i} F$$

(3) Si certains centres ont été modifiés par l'étape (2) retourner en (1) sinon FIN.

3.4.5. - Commentaires sur l'apprentissage

Le procédé de classification automatique en trois étapes que nous avons utilisé est du à la façon dont sont représentées les formes. En effet, celles-ci sont déterminées par des jeux de triplets dont le nombre n'est pas constant. D'autre part, les unités des paramètres des triplets sont différentes. Il est donc hors de question de mesurer les dissemblances entre formes par les distances classiques Euclidienne, Maximini,...

La première étape nous permet de nous affranchir du premier problème en séparant la base d'apprentissage en classes dont les membres sont représentés par le même nombre de paramètres représentant les mêmes phénomènes physiques ; toutes les formes d'une même classe sont supposées être des observations des mêmes vibrations du coeur du réacteur. Les paramètres sont donc maintenant comparables. Ces comparaisons sont effectuées par les distances définies par l'étape 2. La comparaison entre deux formes d'une même classe s'effectue par comparaison deux à deux des triplet-résonances les composant. Les distances inter-résonances proposées sont sans dimension et normalisées de façon à s'affranchir des grandes différences entre les ordres de grandeur des paramètres d'une résonance à une autre.

Ces distances sont pondérées de façon classique par l'inverse de leur valeur moyenne afin d'accorder la même importance discriminante à chacune des résonances. Cette étape permet également la séparation des classes admettant une forte dispersion sur une résonance.

La définition d'une distance inter-forme et du nombre de classes permet l'utilisation d'un algorithme du type de celui des k moyennes qui réalise la classification finale.

Cet apprentissage utilise deux seuils fseuil et Zseuil. Le premier est déterminé de façon automatique, le deuxième de façon expérimentale.

3.5. - La décision

La phase d'apprentissage nous fournit une découpe de la base d'apprentissage en un certain nombre de classes. Il s'agit maintenant d'établir des règles de décision permettant la "reconnaissance" d'une forme inconnue comme membre de telle ou telle classe. La règle retenue s'apparente à un schéma bayesien empirique.

3.5.1. - Choix des fonctions de vraisemblance des classes

Nous définissons la vraisemblance d'une forme F $F\ (R_1, R_2, R_{n_F})$ relativement à une classe ω_i de centre $F^i = (R_1^i, R_2^i, .. R_{n_F i}^i)$ par :

$$\tilde{P}\ (F/\omega_i) = 0 \text{ si } n_F \neq n_{Fi}$$

$$\tilde{P}\ (F/\omega_i) = \prod_{j=1}^{n_{Fi}} \chi_j^i\ (d\ (R_j, R_j^i)) \quad \text{si } n_F = n_{Fi}$$

avec d l'une des distances inter-résonances définies au §. 3.4.2. et χ_j^i la loi du CH1-2 ayant même moyenne et même variance que la moyenne et la variance de la distance inter-résonance estimées empiriquement dans la classe ω_i sur la résonance j.

La loi du χ^2 normalisé est :

$$f(\chi^2) = \frac{1}{2^{\nu/2}\Gamma(\frac{\nu}{2})} \; e^{-\chi^2/2} \; (\chi^2)^{\nu/2-1}$$

avec $\Gamma(\alpha) = \int_0^\infty e^{-x} \, x^{\alpha-1} \, dx$

C'est la loi suivie par la somme des carrés de ν variables gaussiennes normalisées. L'allure de ces densités de probabilité est donnée figure 20. La fonction caractéristique d'une loi du χ^2 générée par la somme des carrés de ν variables gaussiennes centrées d'écart type σ est :

$$\psi(\theta) = (1 - 2\sigma\theta)^{-\nu/2}$$

Les deux premiers moments de cette loi sont :

$$E\left[\chi^2\right] = \frac{\partial \psi}{\partial \theta}(0) = \sigma\nu$$

$$E\left[(\chi^2)^2\right] = \frac{\partial^2 \psi}{\partial \theta^2}(0) = \sigma^2\nu\,(\nu + 2)$$

On en déduit :

$$\nu = \frac{2\,E[\chi^2]^2}{E[(\chi^2)^2] - E[\chi^2]^2}$$

$$\sigma = \frac{E[(\chi^2)^2] - E[\chi^2]^2}{2\,E[\chi^2]}$$

Ce qui donne en remplaçant $E[\chi^2]$ et $E[(\chi^2)^2]$ par les valeurs empiriques $\bar{d}$ et $\bar{d}^2$, les paramètres de notre approximation.

Le choix de ces fonctions de vraisemblance a été dicté par les considérations suivantes :

Une forme ne peut pas être affectée à une classe dont les membres ne sont pas représentés par le même nombre de triplet-résonance. La vraisemblance d'une forme relativement à ces classes doit donc être nulle. L'appartenance d'une forme à une classe est déterminée dans notre étude par les similitudes résonance par résonance entre cette forme et le centre de la classe. Ces similitudes sont mesurées par une distance inter-résonance d. La vraisemblance d'un triplet-résonance par rapport à une résonance de classe peut donc être définie par une approximation de la distribution de probabilité de la distance inter-résonance pour cette résonance et dans cette classe. Les lois du χ^2 réalisent expérimentalement de bonnes approximations. La vraisemblance d'une forme par rapport à une classe est alors définie comme le produit des vraisemblances de chacun des triplet-résonance la composant.

3.5.2. - Décision de reconnaissance

Une forme inconnue F est affectée à la classe la plus vraisemblable à postériori c'est à dire à la classe ω_o telle que :

$$\forall i \quad \tilde{P}(\omega_o/F) > \tilde{P}(\omega_i/F)$$

$$\text{avec } \tilde{P}(\omega_i/F) = \frac{\tilde{P}(F/\omega_i)\,\pi(\omega_i)}{\sum \tilde{P}(F/\omega_j)\,\pi(\omega_j)}$$

où π représente les probabilités à priori de chaque classe.

Ces probabilités sont estimées par les rapports du nombre de formes dans chaque classe au nombre total de formes ayant servi à l'apprentissage.

3.5.3. - Décision d'anormalité

Deux critères d'anormalité ont été expérimentés.

- Vraisemblance

La décision d'anormalité est prise lorsque la classification réalisée au §. 3.5.2. s'avère peu vraisemblable, c'est à dire lorsque la probabilité d'appartenance de la forme à la classe optimale ω^* est inférieure à un seuil dépendant de cette classe et calculé comme suit :

Soit S un paramètre de sensibilité :

$$P\ (F/\omega^*) < S\ (\omega^*) \quad \Longrightarrow \quad \text{Anormalité}$$

$$S\ (\omega^*) = \prod_{j=1}^{n} \chi_j^{\omega^*}\ (d_j^{\omega^*})$$

$$\int_0^{d_j^{\omega^*}} \chi_j^{\omega^*}\ (u)\ du = 1 - S \qquad \forall j.$$

- Ecart maximum

La décision d'anormalité est prise si l'une des résonances de la forme s'écarte trop de la résonance correspondante de la classe identifiée.

$$\exists j \quad d_j^{\omega}\ (R_j, R_j^{\omega^*}) > d_j^{\omega^*} \quad \Longrightarrow \quad \text{Anormalité}$$

$$\int_0^{d_j^{\omega^*}} \chi_j^{\omega^*}\ (u)\ du = 1 - S \qquad \forall j$$

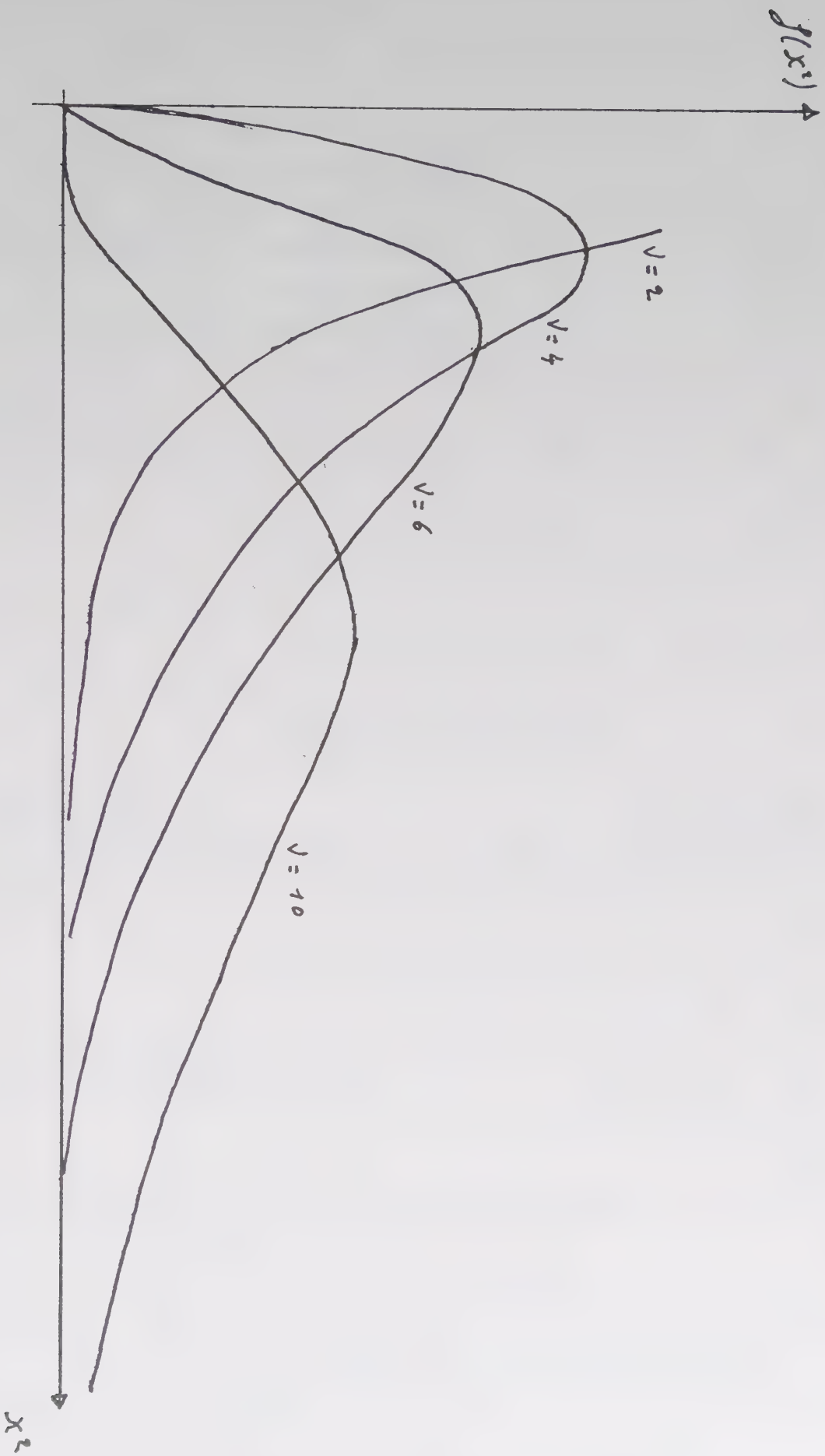


Fig 20: Allure des lois du CHI-2

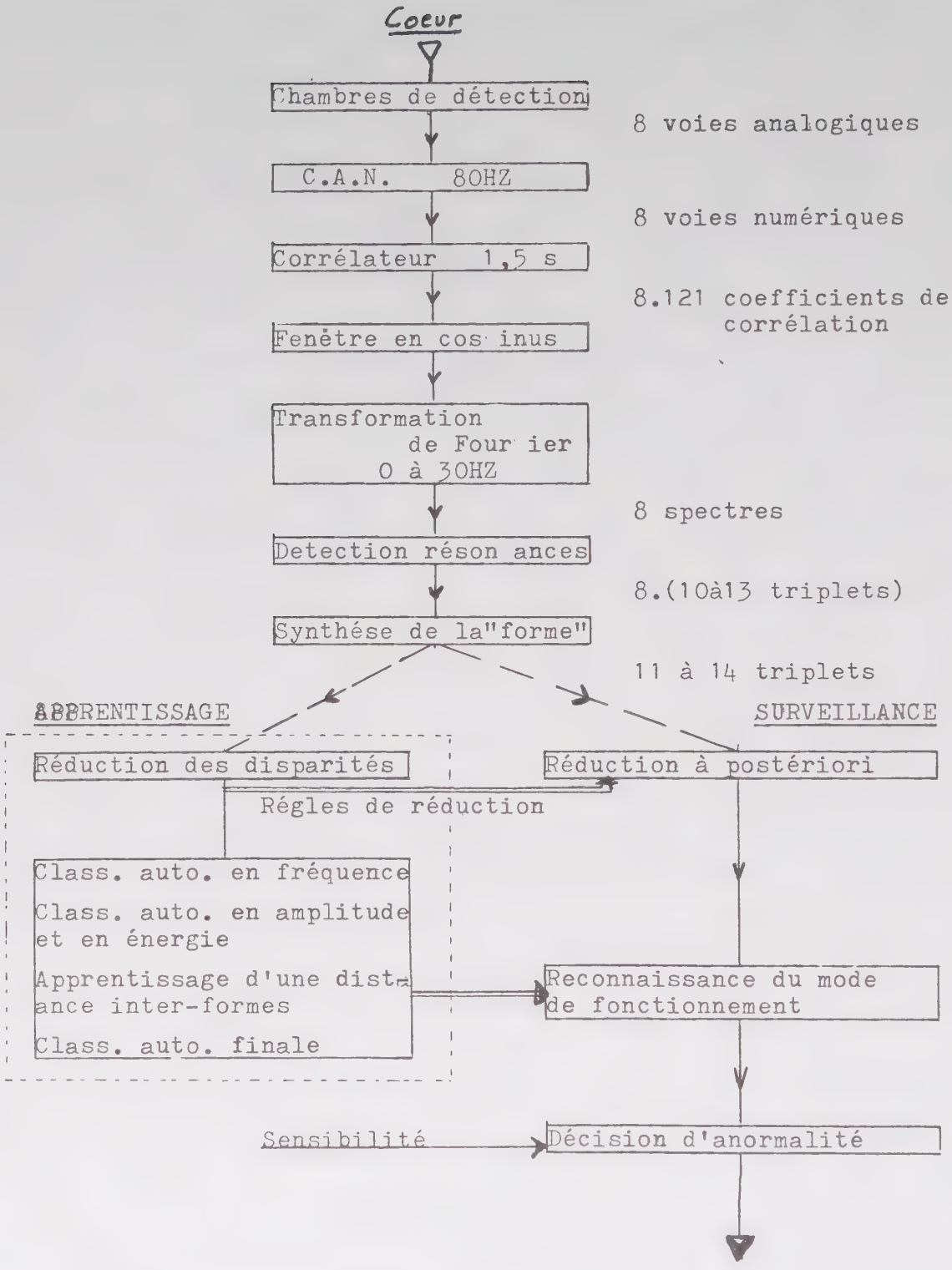


Fig 21 : Méthode de surveillance

3.6. - Application informatique

La principe général de la méthode de surveillance proposée est présenté figure 21. L'application réalisée se compose de deux programmes : APREN et RECON. Le premier réalisé à partir d'une base de paramètres extraits l'apprentissage des règles de réduction à postériori, de la classification et des paramètres nécessaires à la décision. Son travail est symbolisé par la zône entourée en pointillés de la figure 21. Le deuxième réalise, pour une forme quelconque, l'extraction des paramètres et la reconnaissance.

3.6.1. - Le programme APREN

Vu de l'extérieur son fonctionnement est le suivant :

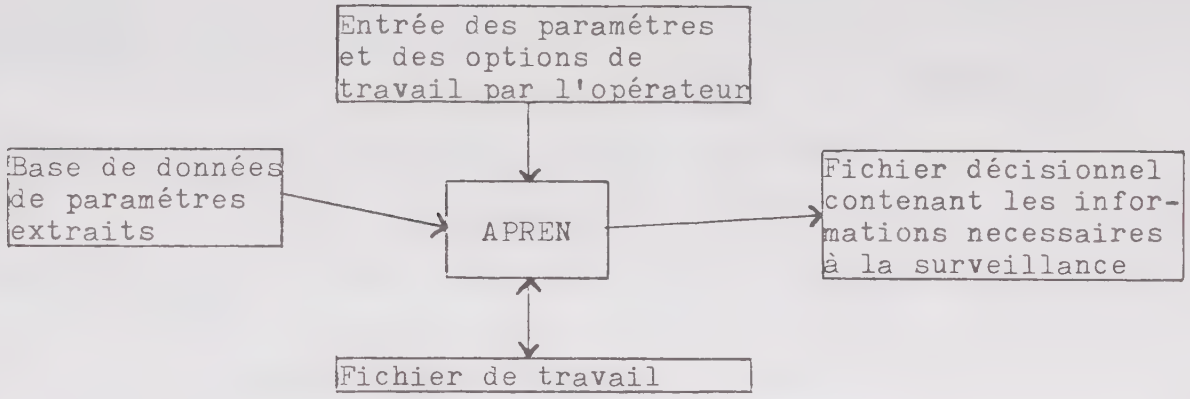


Fig 22: Le programme "APREN"

L'opérateur initialise l'apprentissage en fournissant les renseignements suivants :

- Nom du fichier disque contenant les jeux de paramètres correspondant aux formes à étudier.
- Seuil pour classification automatique en fréquence Fseuil (1 Hz).

- Seuil pour classification automatique en distance Zseuil (1,5)
- Seuil pour l'algorithme réducteur d'entropie Pseuil (0,3)
- Nom du fichier qui contiendra les résultats de l'apprentissage (fichier décisionnel).
- Nom du fichier de travail.

Le programme réalise ensuite l'apprentissage de la réduction à postériori, transforme la base d'apprentissage et la mémorise dans le fichier de travail. La classification automatique est ensuite effectuée. Des options de visualisation d'histogrammes, de probabilité... permettent de suivre les différentes étapes et d'étudier l'influence de tel ou tel paramètre.

Le fichier décisionnel est engendré et contient les informations suivantes :

- Nombre de classes.
- Centre de chaque classe.
- Moyenne et variance par classe et par résonance de la distance inter-résonance.
- Vecteurs de transformation de l'algorithme de réduction à postériori.

Le programme de projection non linéaire exposé au §. 3.6.3. y ajoute dans une phase ultérieure les coordonnées des points d'une représentation plane du nuage des centres de classe.

3.6.2. - Le programme "RECON"

Vu de l'extérieur son fonctionnement est le suivant :

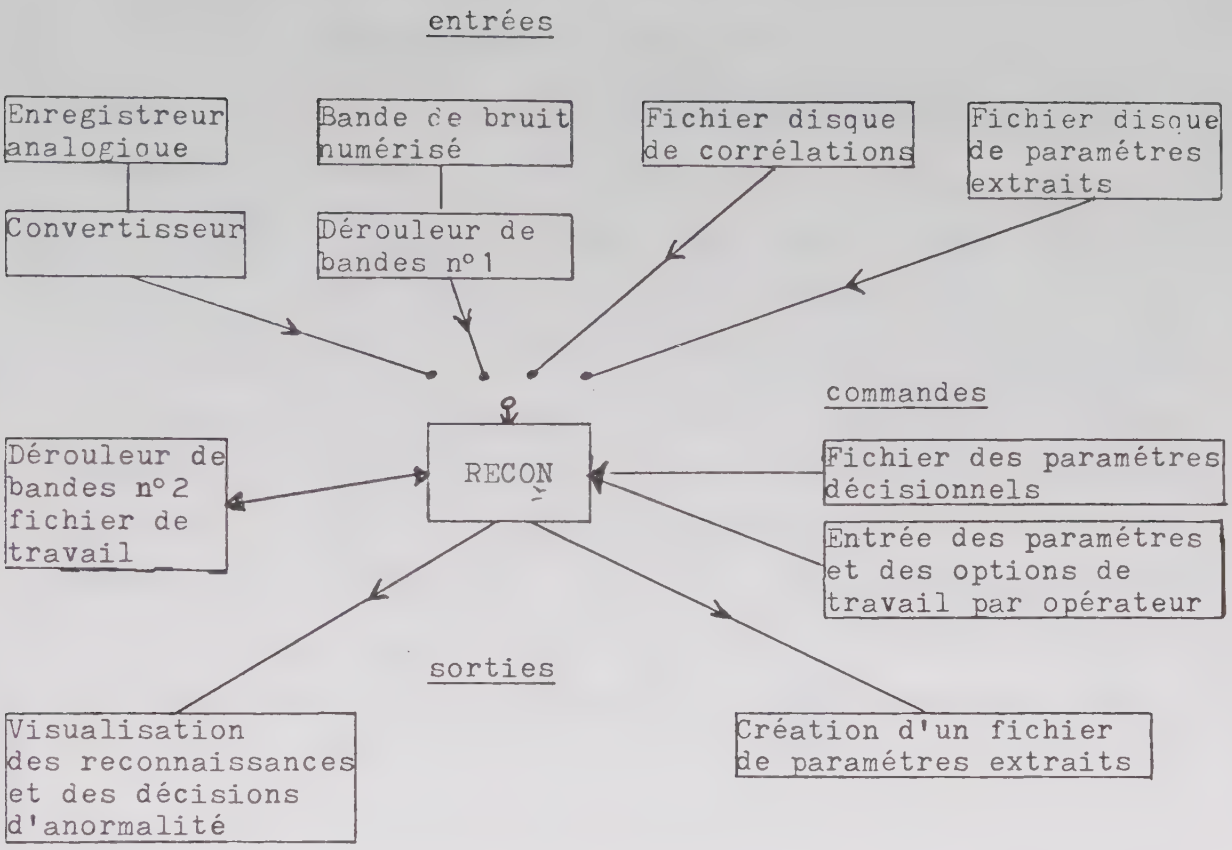


Fig 23: Le programme "RECON"

L'opérateur initialise la reconnaissance en fournissant les renseignements suivants :

- Nom du fichier décisionnel généré par l'apprentissage.
- Paramètre de sensibilité.
- Type des données à reconnaitre :
  . Bande analogique
  . Bande numérique
  . Fichier de corrélations
  . Fichier de paramètres extraits.

En option, à mesure que la reconnaissance est poursuivie un nouveau fichier de paramètres extraits est engendré.

3.6.3. - Visualisations des classes

- Spectres reconstitués

La réduction de la forme initiale à un jeu de triplets, qui nous permet un traitement simplifié présente l'inconvénient de ne pas être facilement visualisable, c'est pourquoi nous avons conçu un programme permettant à des fins de visualisation, la synthèse à partir d'un jeu de paramètres extraits d'un spectre fictif. Cette synthèse s'effectue en supposant chacune des résonances indépendantes et en sommant leurs spectres reconstitués comme au §. 3.4.2. Si le jeu de paramètres extraits est $((f_1\ a_1\ e_1 \ldots (f_P,\ a_P\ e_P))$ le spectre fictif aura l'équation suivante :

$$S(x) = \sum_{i=1}^{P} \frac{u_P\ a_P\ f_P^2}{(x^2-f_P^2)^2 + u_P\ x^2} \qquad u_P = \alpha\left(\frac{e_P}{a_P}\right)^2$$

La figure 24 montre en traits fins les 8 spectres correspondant à une forme et en traits larges le spectre fictif généré à l'aide du jeu de paramètres extraits.

Lors d'une phase de reconnaissance les spectres fictifs des centres des classes sont visualisés sur la partie gauche de l'écran. Ces spectres sont repérés par l'indice de leur classe.

- Projection non linéaire

Chaque classe étant supposée représentée par le spectre reconstitué de son centre, on peut mesurer les "écarts" entre classes par la norme L^2 entre logarithmes de ces spectres. Une représentation plane de ces écarts peut être obtenue par projection non linéaire respectant un critère de proximité. Nous avons choisi comme critère celui de [KRUSHAL].

$$K = \sum_{j>i} \sum_{i=1}^{Nclasse} \left(\sqrt{(x_i-x_j)^2 + (y_i-y_j)^2} - D_{ij}\right)^2$$

avec (x_i, y_i) les coordonnées de la représentation plane de la classe ω_i.

et D_{ij} la distance entre les classes ω_i et ω_j dans l'espace de départ. La minimisation de ce critère est réalisée par l'algorithme suivant :

(1) Déterminer la plus grande distance inter-classe dans l'espace de départ : DMAX.

(2) Soit N la racine carrée du nombre de classes. Initialiser les points (x_i, y_i) aux noeuds d'un réseau carré N+1 x N+1 de maille élémentaire DMAX/(n+1)/1,5
Poser P = DMAX/(N+1)/1.5/10
Poser $K_o = 8$

(3) Considérer chaque point M
Calculer la valeur du critère obtenue en remplaçant ce point M par un point M' obtenu par translation de vecteur $\vec{u}$: $(\cos\theta\ \vec{i} + \sin\theta\ \vec{j})$xP
et ce pour les valeurs : $\frac{2k\pi}{n}$ k = 1, n de θ.
Effectuer les déplacements diminuant le critère.

(4) Si aucun point n'a été déplacé par l'étape (3) aller en (5) .
Sinon retourner en (3) .

(5) Si $K_o = 32$ aller en (6) .
Sinon $K_o = K_o \times 2$; P = P/2 aller en (3) .

(6) Mémoriser la valeur du critère obtenu. Eventuellement recommencer le processus avec une initialisation différente si le résultat n'est pas satisfaisant.

(7) Mémoriser la configuration finale.

Lors de la phase de reconnaissance, la représentation plane des classes est visualisée sur la partie droite de l'écran. Elle permet de suivre par un marquage automatique l'évolution des affectations ainsi que les décisions d'anormalité.

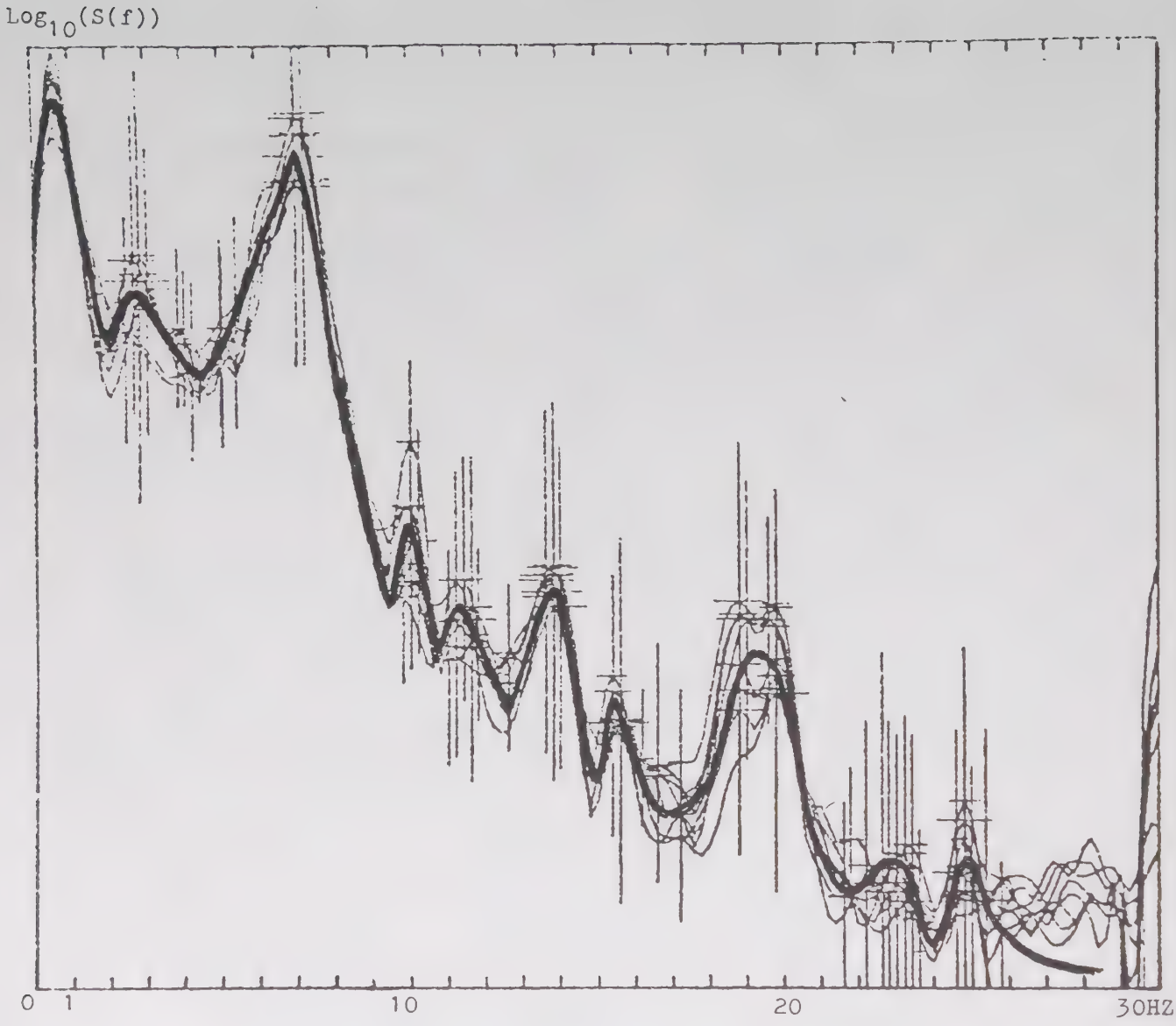


Fig 24: Spectres réels et spectre fictif obtenu

3.7. - Résultats expérimentaux

3.7.1. - Paramètres utilisés

La durée T_f de chaque forme a été fixée à 1 mm. Cette durée est la plus faible permettant une estimation avec une variance inférieure à 10^{-3} des coefficients de corrélations. Nous avons choisi cette durée minimale de façon à pouvoir considérer un nombre important de formes différentes. Dans le cas d'une application en ligne il ne serait demandé qu'une décision toutes les dix ou vingt minutes.

La corrélation est estimée sur une fenêtre de 1,5 s. ce qui à la fréquence de 80 Hz correspond à 121 coefficients. Ce nombre est suffisant pour la définition en fréquence des spectres désirée. Le temps de coupure t_c de la fenêtre en cosinus surélevé a été choisi à 0,4 s. Les spectres sont calculés de 0 à 30 Hz sur 151 points.

La courbe ; nombre moyen de fréquences de résonance par forme en fonction de la valeur du seuil de réduction est donnée figure 25. La valeur du Fseuil retenue est 1 Hz, elle correspond au début de la zône stationnaire de cette courbe.

La figure 26 illustre le fonctionnement de l'algorithme de réduction. Les huit lignes du bas montrent les positions de chacune des résonances sur les huit voies. Les traits horizontaux symbolisent les associations réalisées, leur ordonnée est l'écart en fréquence entre centres associés.

La courbe : écart moyen en fréquence d'une forme au centre de sa classe d'appartenance après l'étape du 3.4.1 est donnée figure 27. La valeur du fseuil choisie est celle qui minimise cette courbe soit 0,4 Hz. Le paramètre Zseuil a été expérimentalement fixé à 1,5. Le critère d'arrêt retenu de l'algorithme des K moyennes est la convergence stricte.

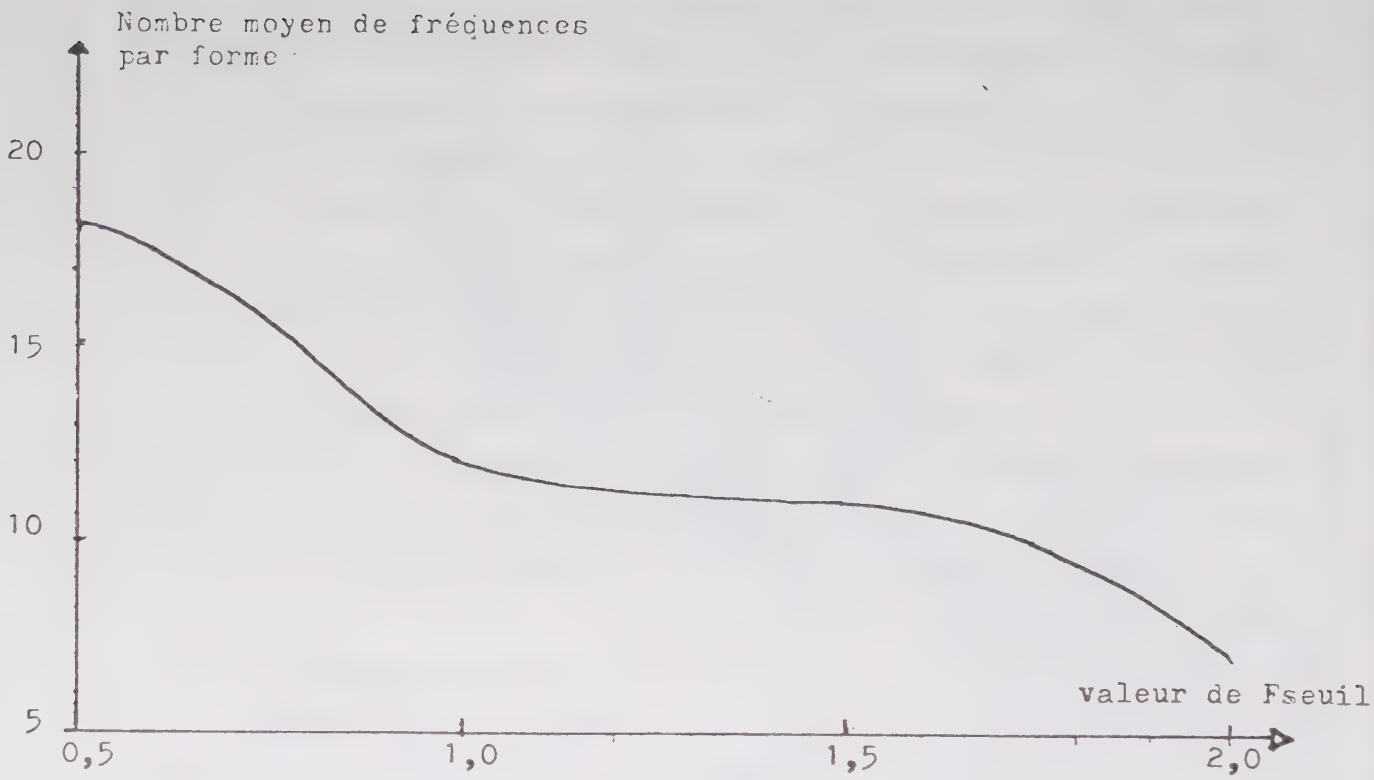
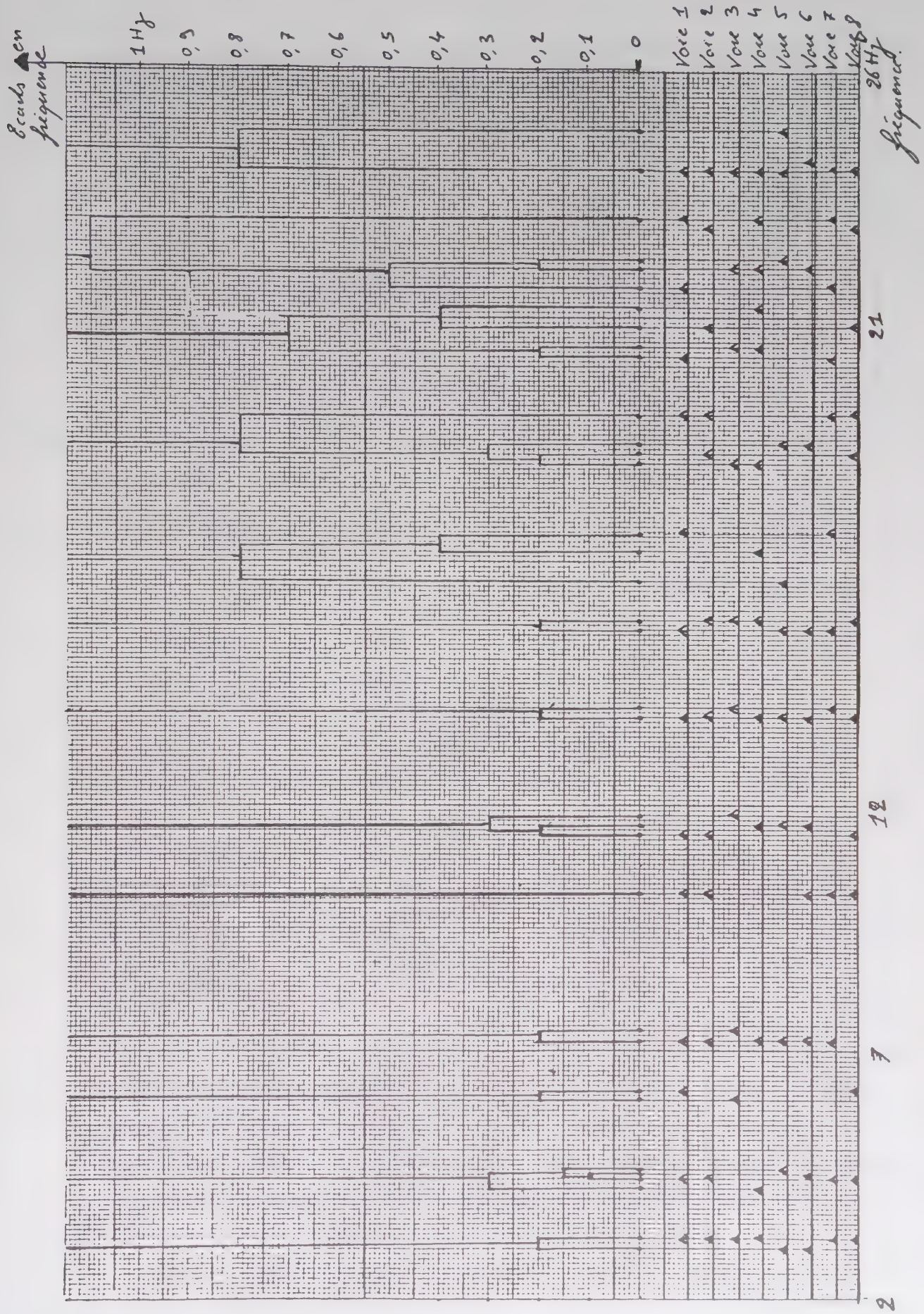
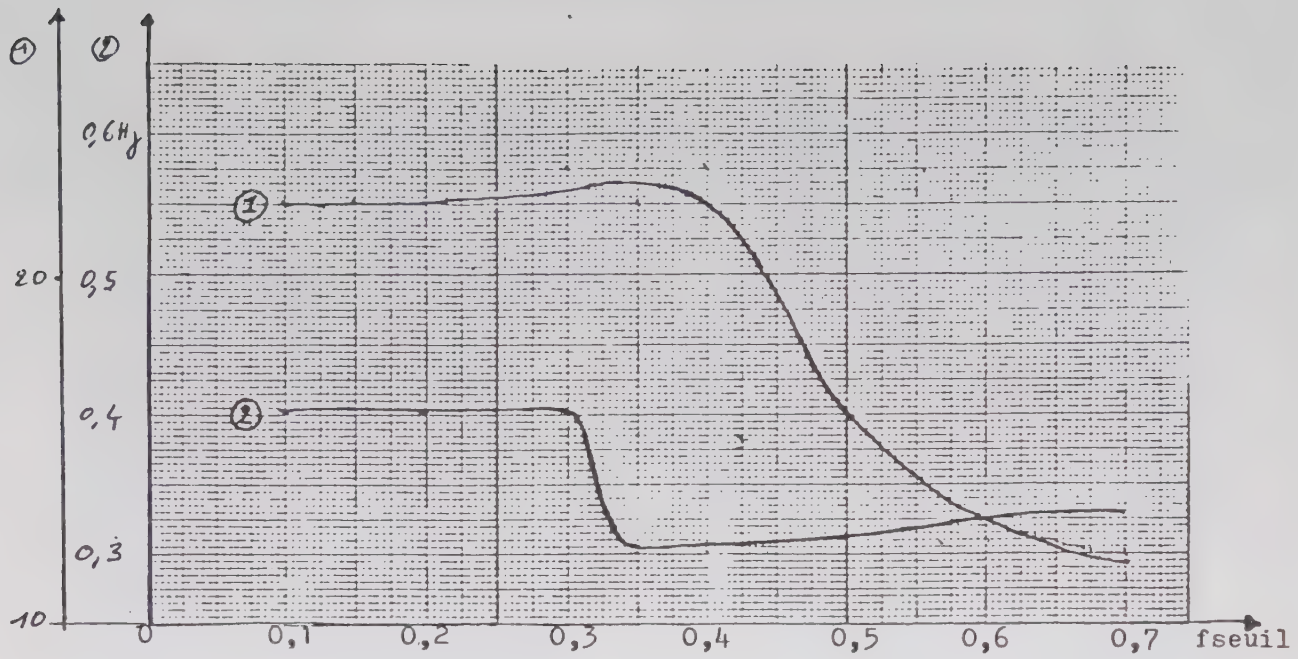


Fig 25: Nombre moyen de fréquences par forme en fonction de la valeur du paramétre Fseuil

Fig 26: Illustration de la réduction par classification automatique





1: Nombre de groupes
2: Ecart moyen en fréquence entre une forme et le centre de son groupe

Fig 27: Détermination de la valeur du fseuil optimal

3.7.2. - Résultats généraux et utilisation

Un résultat important à vérifier lorsque l'on fait appel à un apprentissage non supervisé est la stabilité des classes trouvées.

La figure 28 montre l'évolution du nombre de classes en fonction du nombre de formes d'apprentissage. Il semble que cette courbe tend vers une asymptote horizontale. D'autre part, l'étude des classes déterminées montre qu'il n'existe pas de corrélation entre les classes et la date des mesures. Dans chaque classe on trouve des formes issues de mesures de dates différentes. La méthode de classification automatique que nous avons utilisée n'assure pas une partitionnement exact de la base d'apprentissage. Il existe des formes non représentées. Le nombre de ces formes, influant sur la probabilité minimum de fausse alarme, doit être le plus faible possible. La courbe de la figure 29 représente la proportion de formes non représentées en fonction du nombre de formes de la base d'apprentissage.

Lors de la phase de reconnaissance, les seules erreurs de classification relevées ont été causées par le jeu des probabilités à priori de la décision baysienne.

Une classification au maximum de vraisemblance n'eût causé aucune erreur. Ceci laisse supposer une forte séparabilité de la partition obtenue.

Importance de la base d'apprentissage

La notion d'anormalité est définie par apparition aux données de la base d'apprentissage. Celle-ci doit donc être un résumé exhaustif de ce que l'on considère comme normal. Cela suppose un grand nombre d'enregistrements correspondant à toutes les situations de fonctionnement que l'on veut surveiller. Il ne doit pas y avoir, dans cette base, d'enregistrements anormaux ne serait-ce que par des accidents de mesure. Car cela entérinerait une forme d'anormalité. Dans l'approche que nous avons adopté aucune supposition quant à la nature des anormalités possibles n'a été émise. Le choix des distances et des algorithmes a été effectué sur des considérations générales et ne privilégie pas tel ou tel aspect des formes. Il est certain qu'une définition plus précise des types d'anormalité que l'on se proposerait de détecter permettrait des choix plus orientés. Il est toutefois possible sans changer la méthode de s'intéresser en plus à certaines anormalités connues. Il suffit pour ce faire d'introduire dans la base de données des enregistrements anormaux identifiés et lors de la phase de reconnaissance de prendre la décision d'anormalité lorsque la forme inconnue est affectée à une classe de formes anormales. Un diagnostic peut donc ainsi être réalisé. La consultation d'un journal de bord du réacteur nucléaire devrait permettre de donner une signification physique aux différentes classes et une interprétation du fait d'assigner une forme à une classe donnée. Un contrôle du fonctionnement du réacteur serait alors possible.

Dans notre application nous n'avons utilisé que les chambres de détection neutronique. Tout autre capteur indépendamment de la nature du signal qu'il fournit (discret, binéaire, bruit...) peut être pris en compte par une approche "reconnaissance de formes". Il est même conçevable que tous les capteurs d'un réacteur soient considérés. Une réaction de l'apprentissage sur la sélection de ceux-ci basée sur l'optimisation de critères d'efficacité permettrait le choix, en ligne des capteurs à considérer suivant le type d'anormalité ou de fonctionnement reconnu.

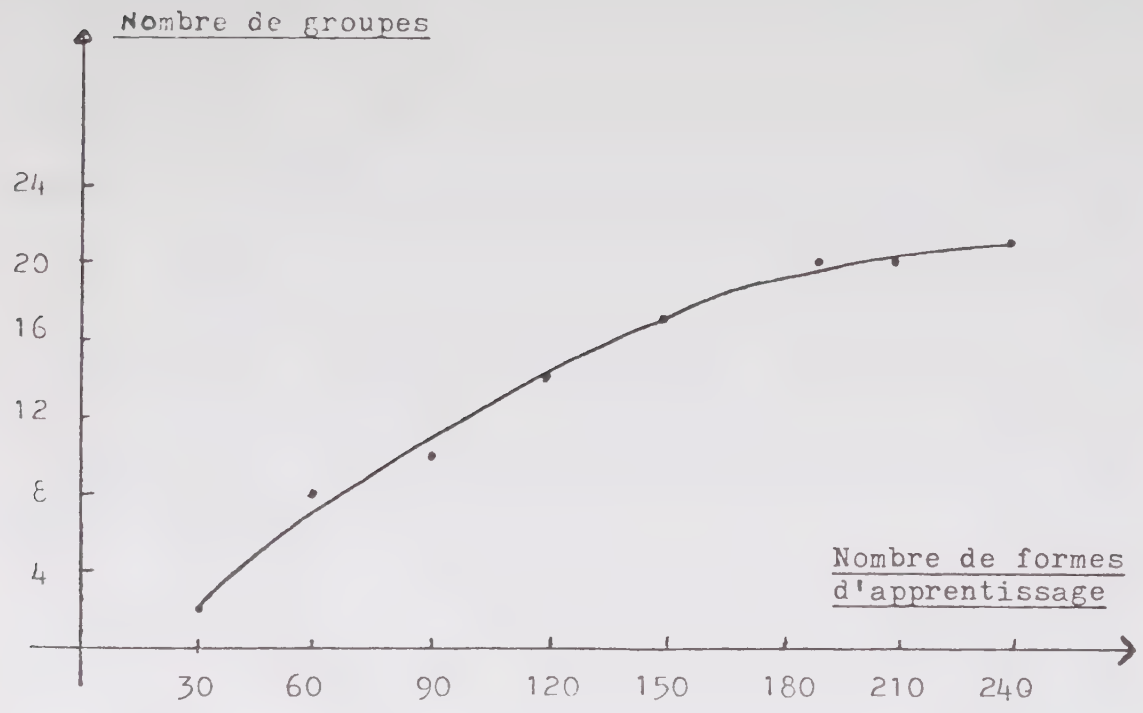


Fig 28: Evolution du nombre de classes en fonction de la taille de la base d'apprentissage

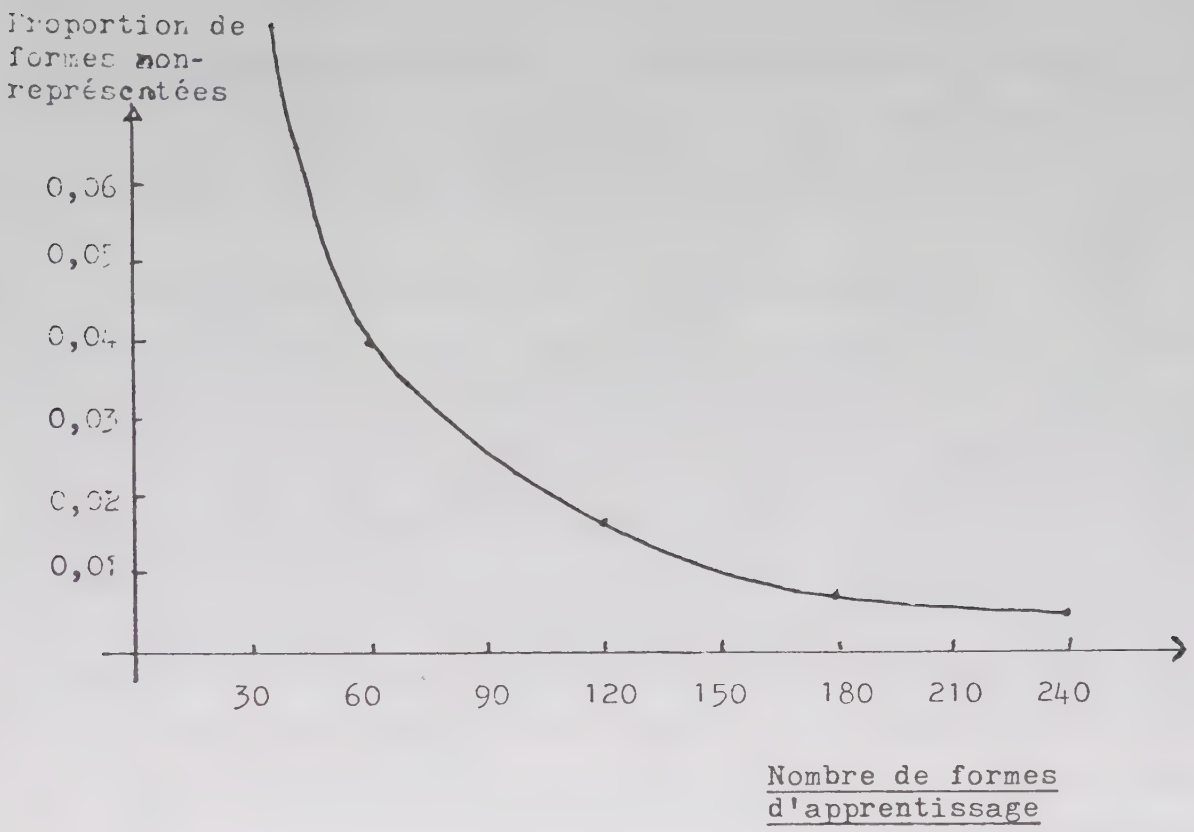


Fig 29: Evolution du nombre de formes non représentées en fonction de la taille de la base d'apprentissage

3.7.3. - Probabilité de fausse alarme

Une détermination empirique des probabilités de fausse alarme a été réalisée de la façon suivante :

Apprentissage sur 120 formes d'indices pairs, reconnaissance sur 120 formes d'indices impairs, puis apprentissage sur 120 formes d'indices impairs et reconnaissance sur 120 formes d'indice pair. La probabilité de fausse alarme étant alors estimée par le rapport du nombre d'alarmes sur 240. Cette expérience a été reconduite pour différentes valeurs du paramètre de sensibilité et pour les différentes versions proposées de la méthode.

[1] Distance inter-résonance d. - Décision d'anormalité "vraisemblance".

[2] Distance inter-résonance d, - Décision d'anormalité : écart maximum

[3] Distance inter-résonance d', - Décision d'anormalité : vraisemblance

[4] Distance inter-résonance d', - Décision d'anormalité : écart maximum.

Les figures 30, 31, 32 et 33 montrent respectivement pour chacune de ces versions les courbes de la variation des probabilités de fausse alarme en fonction du paramètre de sensibilité. Les courbes numérotées |1| correspondent à des reconnaissances sur des formes non issues de la base d'apprentissage. Les courbes numérotées |2| sont obtenues par reconnaissance sur les formes ayant servies à l'apprentissage. Les courbes numérotées |3| et |4| sont identiques aux 1 et 2 mais la décision d'anormalité y est pondérée par la règle : une anormalité est effective si dans 3 formes successives dans le temps deux anormalités ont été détectées.

3.7.4. - Détection d'une fréquence parasite

Afin de déterminer l'aptitude de cette méthode de surveillance à détecter une fréquence parasite, nous avons procédé aux expériences suivantes :

Apprentissage sur 120 formes, reconnaissance sur 30 autres formes perturbées par une fréquence parasite (ajout d'une sinusoïde pure sur les 8 voies). La décision d'anormalité y est pondérée par la règle exposée au §.3.7.3. Les courbes nombre d'anormalités détectées en fonction de la fréquence ont été tracées.

Les figures 34, 35, 36 et 37 correspondent aux versions 1, 2, 3 et 4 et à un rapport Energie bruit neutronique/Energie fréquence parasite de 14 dB.

Les figures 38, 39, 40 et 41 correspondent aux versions 1, 2, 3 et 4 et à un rapport signal/perturbation de 34 dB. En règle générale, ces courbes imposent le commentaire suivant : une fréquence parasite située sur une résonance est plus difficile à détecter.

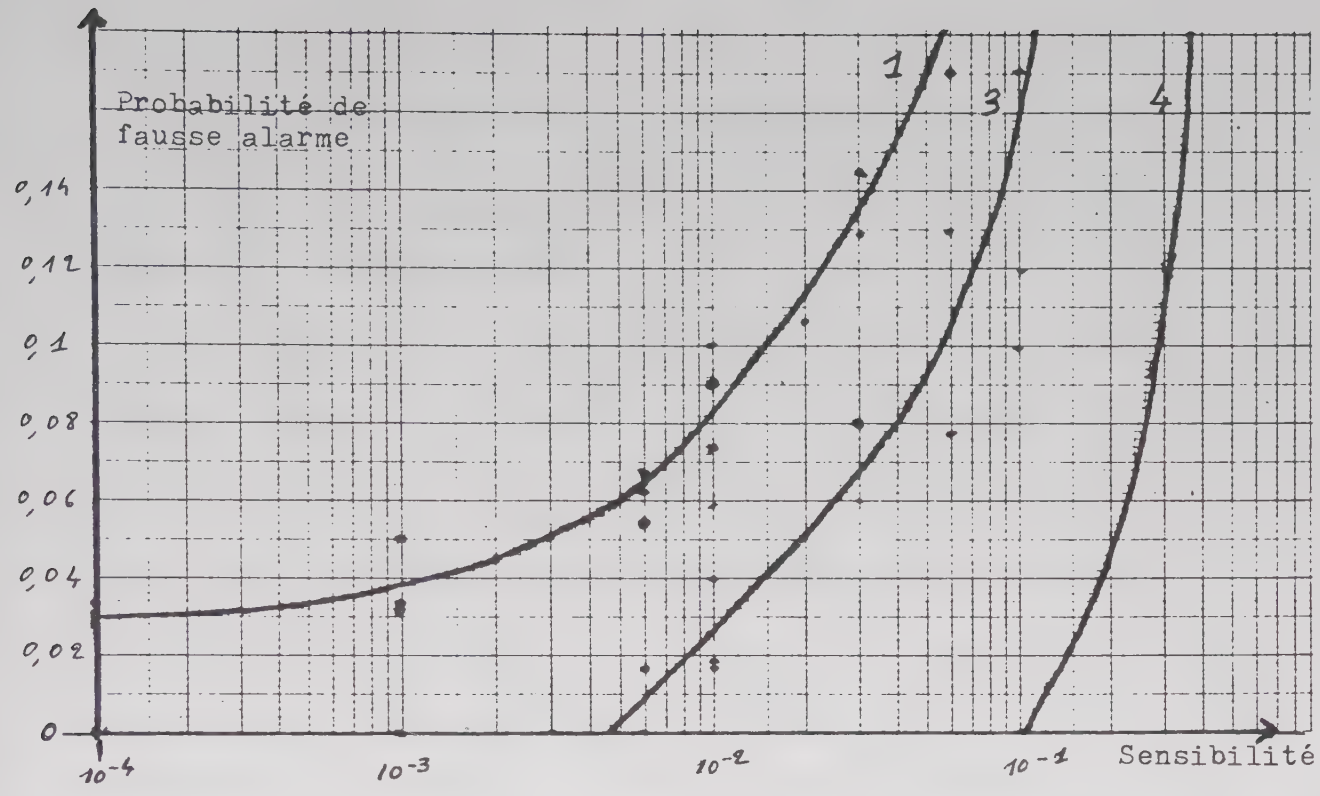


Fig 30 : Probabilités de fausse alarme version n° 1

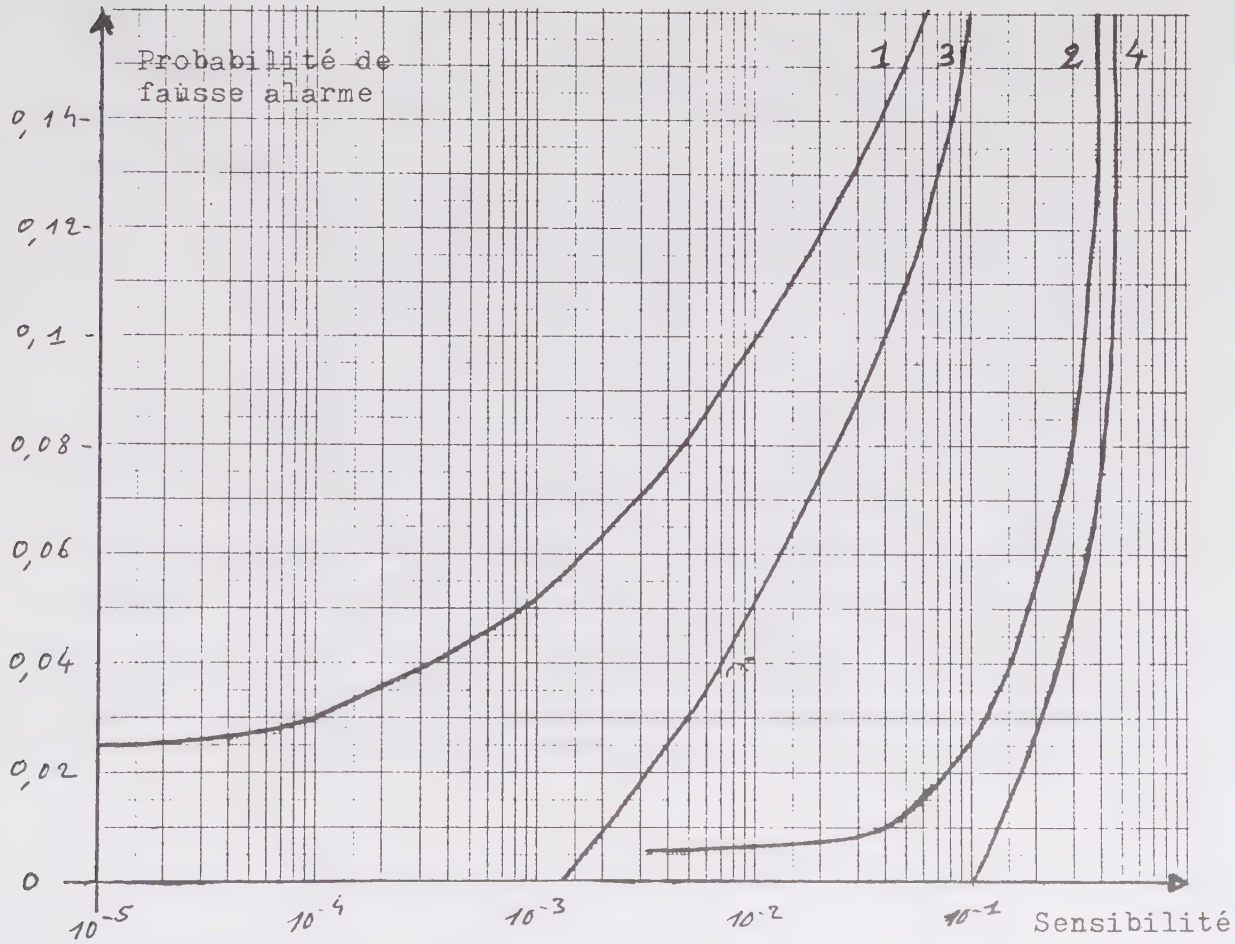


Fig 31 : Probabilités de fausse alarme version n° 2

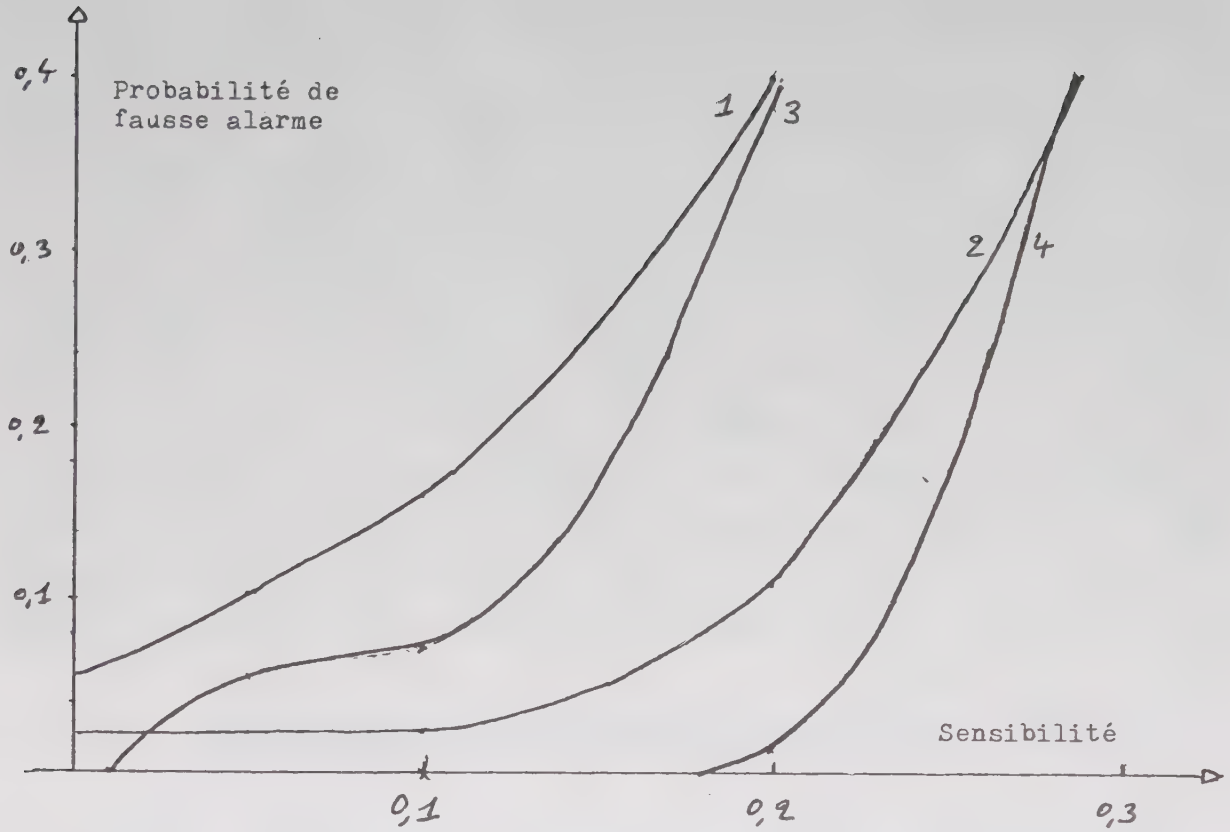


Fig 32: Probabilités de fausse alarme version n°3

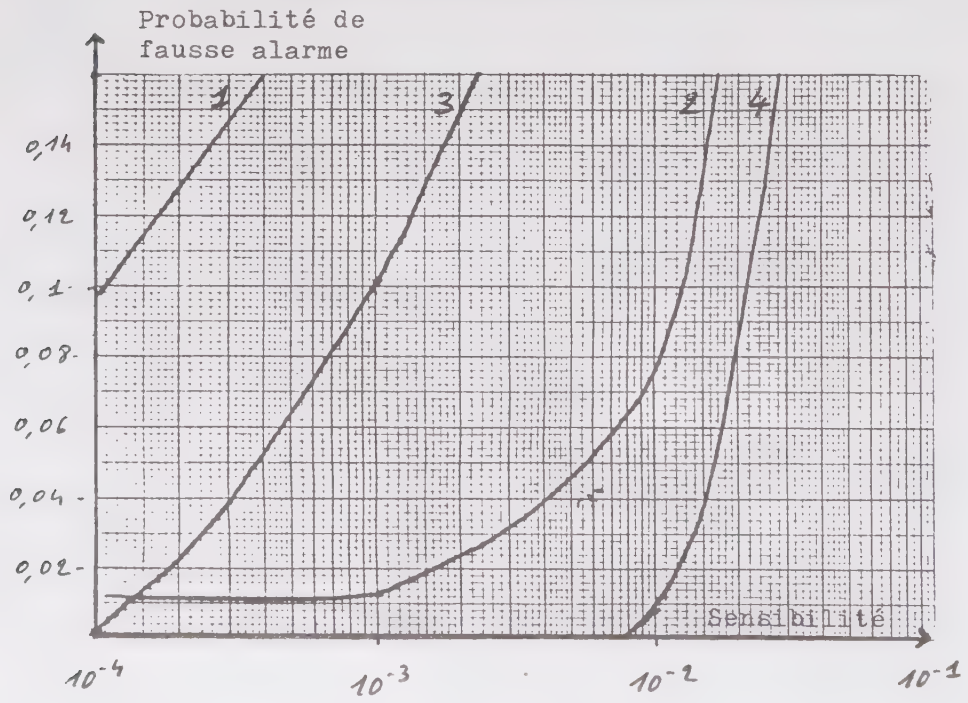


Fig 33: Probabilités de fausse alarme version n°4

Dans chacune de ces expériences la sensibilité a été ajustée à la plus grande valeur assurant une probabilité de fausse alarme nulle sur les 120 formes n'ayant pas servies à l'apprentissage.

Aux vues de ces courbes, les quatre versions semblent devoir se classer dans l'ordre de qualité décroissante suivant : 4→2→3→1. En fait, les résultats obtenus sont comparables à la précision des approximations près. Par contre la version n° 3 présente l'avantage d'admettre une dépendance probabilité de fausse alarme/sensibilité beaucoup plus facile à optimiser.

Nous avons donc restreint la suite de notre étude expérimentale à cette version, c'est à dire : distance inter-résonance d', décision d'anormalité sur critère de vraisemblance.

La figure 42 montre la courbe : Rapport bruit signal minimum nécessaire pour obtenir une probabilité d'alerte de 0,5 en ayant une probabilité de fausse alarme inférieure à la précision de l'expérience soit 1/120 pour une fréquence parasite variant entre 0 et 30 Hz.

3.7.5. - Bruit blanc parasite

Les figures 43 et 44 montrent le nombre d'anomalies détectées lorsqu'un bruit blanc est introduit sur une voie (fig.43) ou sur deux voies (fig.44). Les courbes repérées par des points correspondent aux décisions d'anormalité "vraies", celles repérées par des croix, à la règle de pondération : une anormalité est effective si dans trois formes consécutives deux sont anormales. Notons que dans ce dernier cas sur 30 formes seules 28 décisions peuvent être prises.

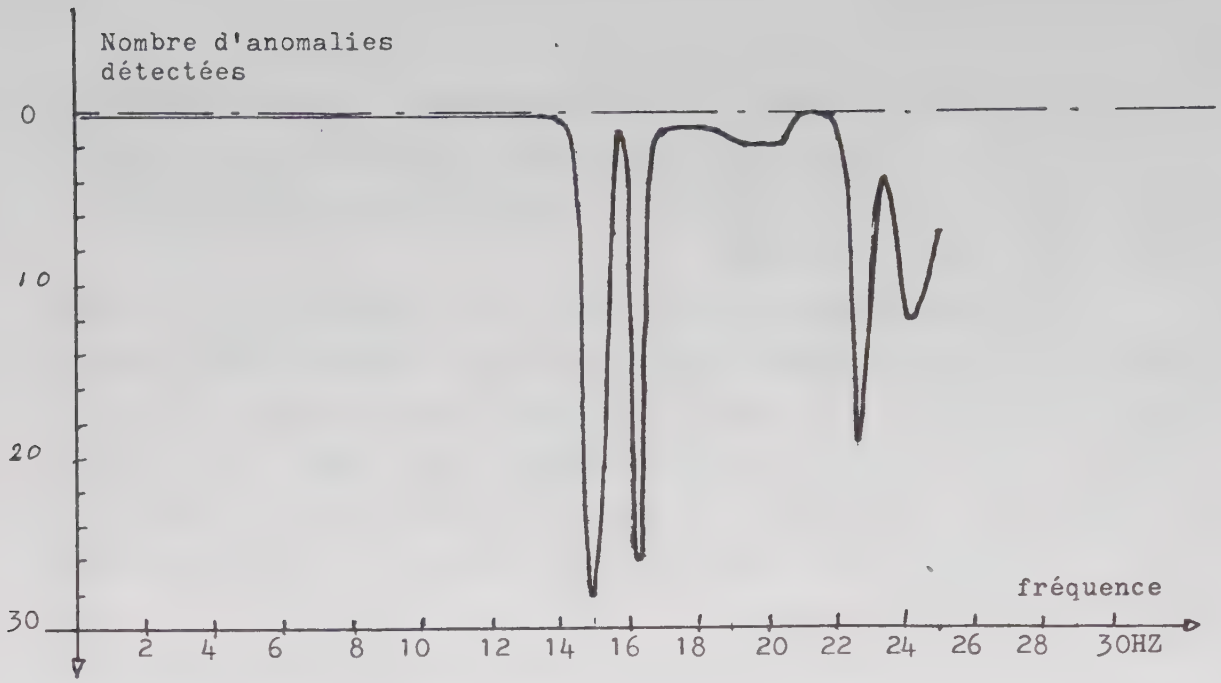


Fig 34: Détection d'une fréquence parasite rapport signal/bruit 14db version n°1

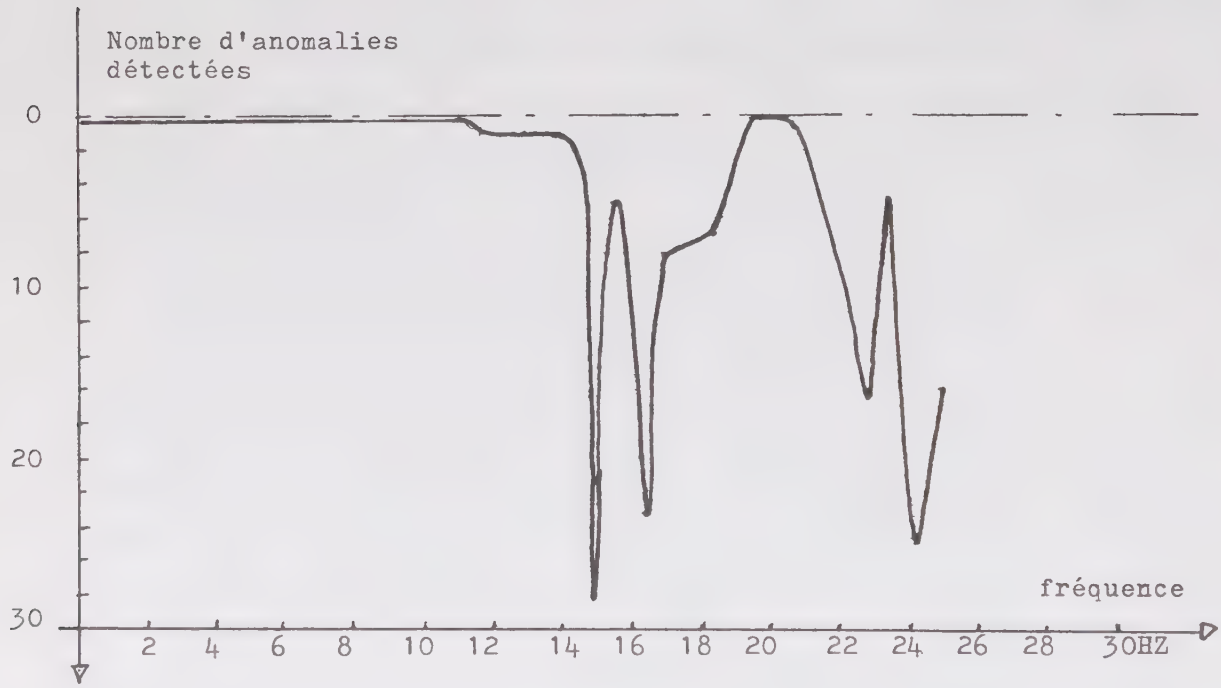


Fig 35: Détection d'une fréquence parasite rapport signal/bruit 14db version n°2

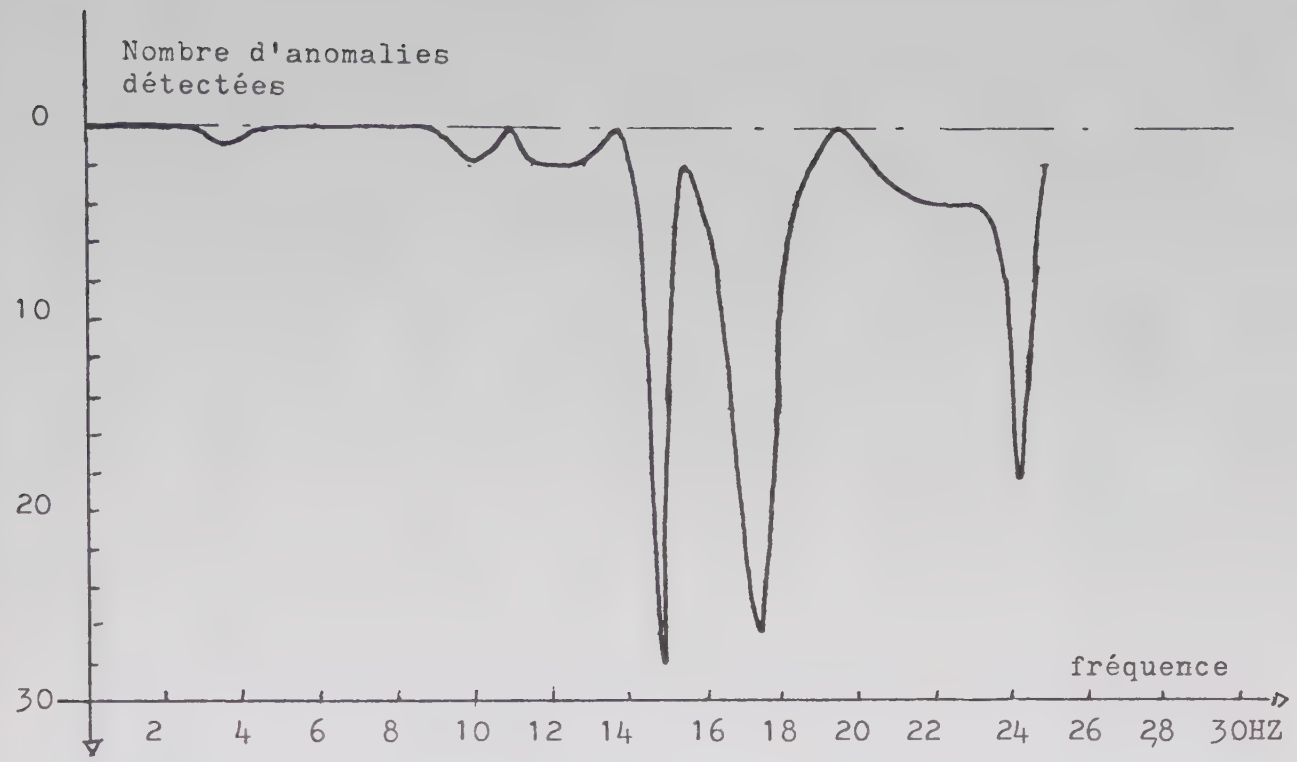


Fig 36: Détection d'une fréquence parasite
rapport signal/bruit 14db version n°3

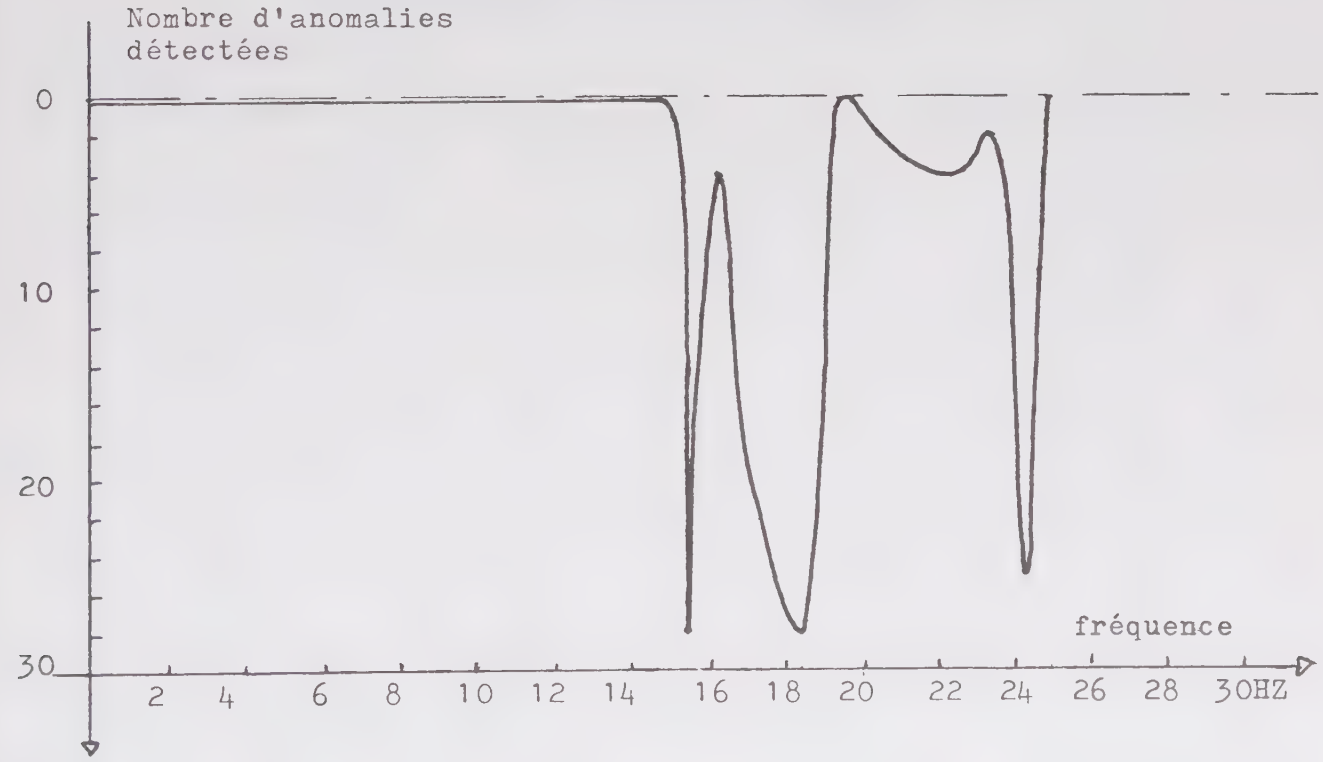


Fig 37: Détection d'une fréquence parasite
rapport signal/bruit 14db version n°4

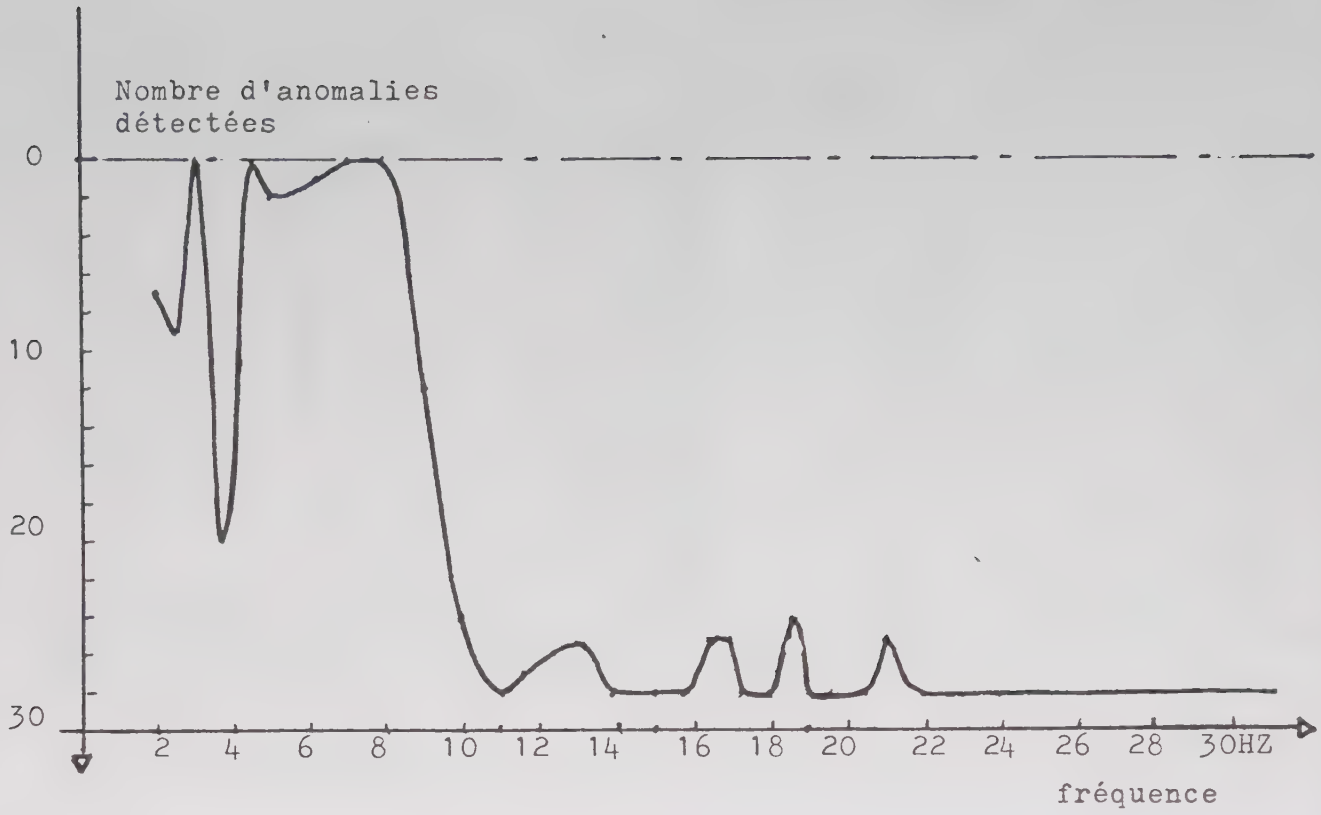


Fig 38: Détection d'une fréquence parasite
rapport signal/bruit 34db version n°1

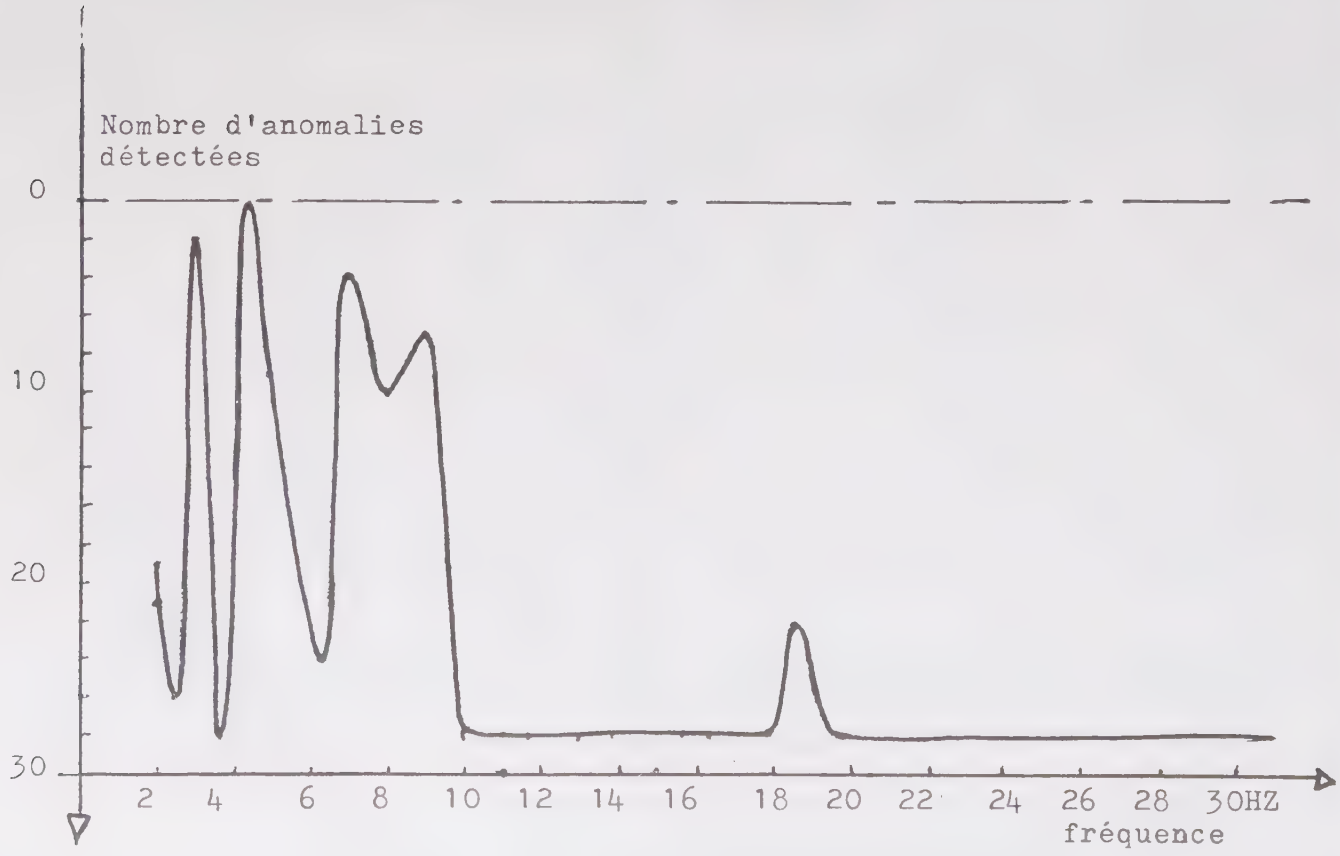


Fig 39: Détection d'une fréquence parasite
rapport signal/bruit 34db version n°2

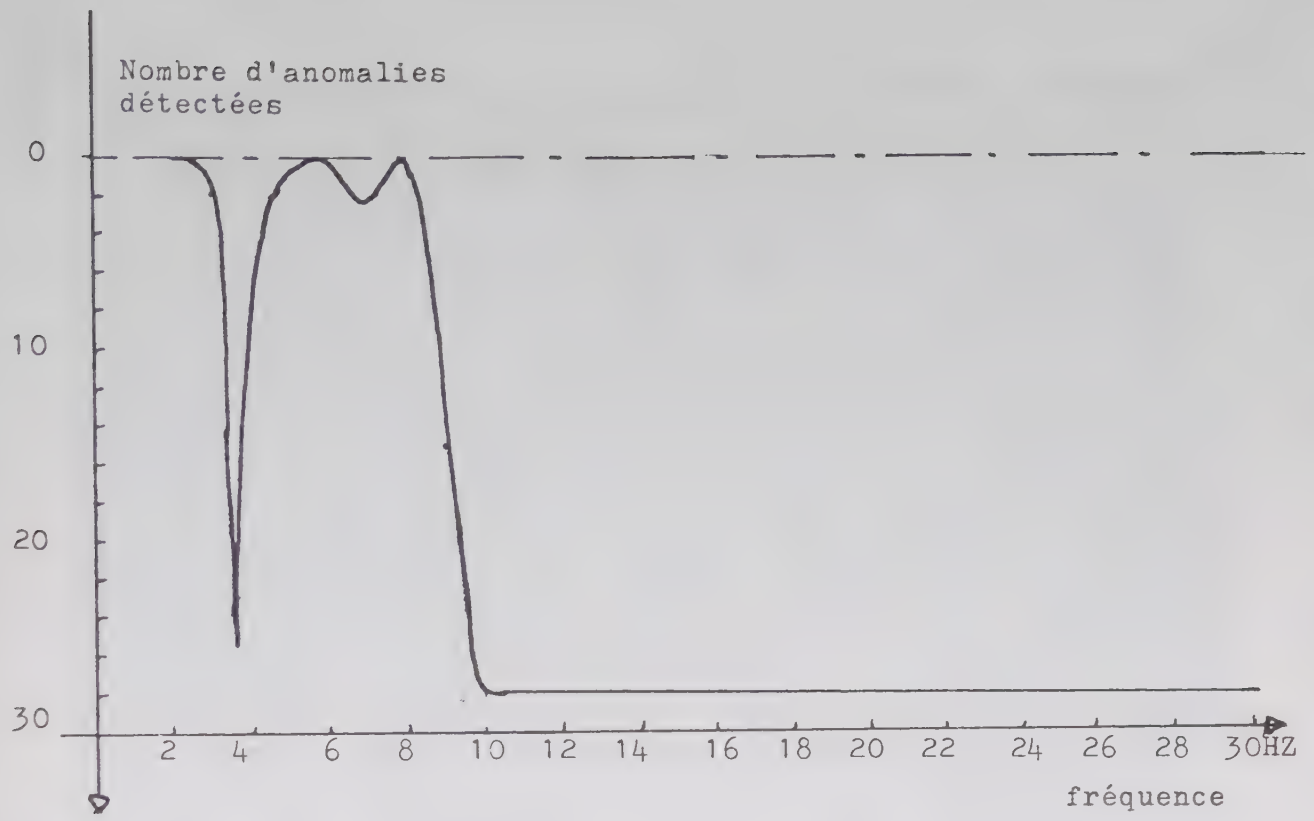


Fig 40: Détection d'une fréquence parasite
rapport signal/bruit 34db version n°3

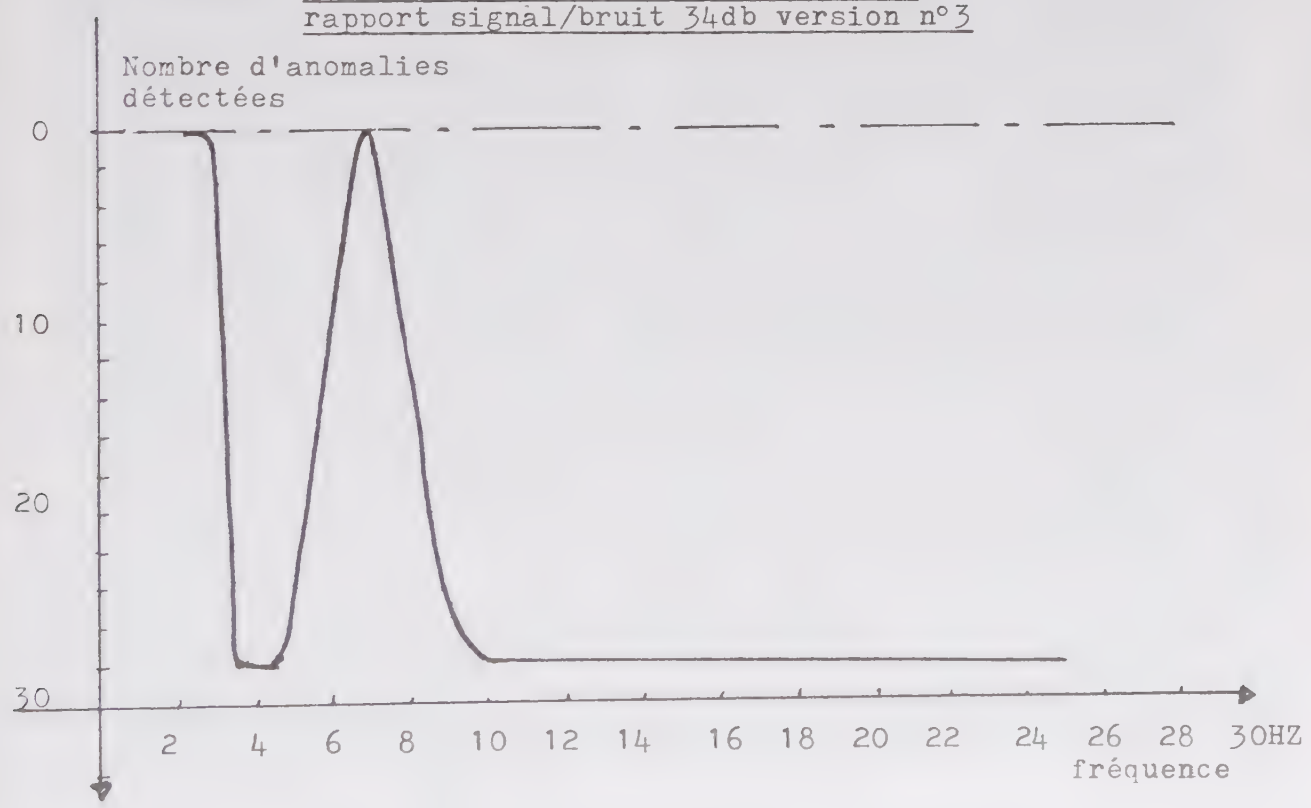


Fig 41: Détection d'une fréquence parasite
rapport signal/bruit 34db version n°4

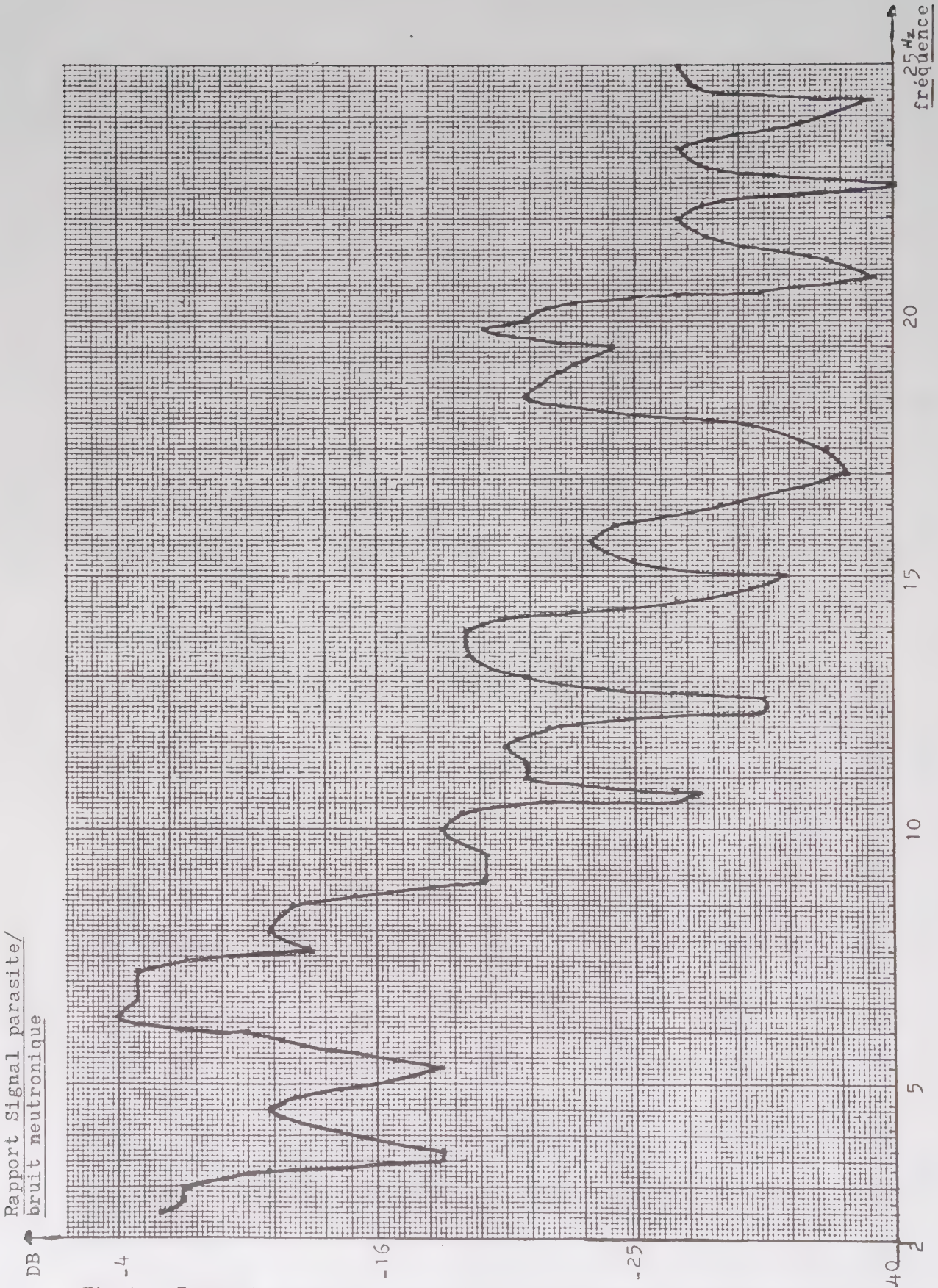


Fig 42 : Rapport signal parasite/bruit neutronique minimum pour obtenir une probabilite de détection de 0.5

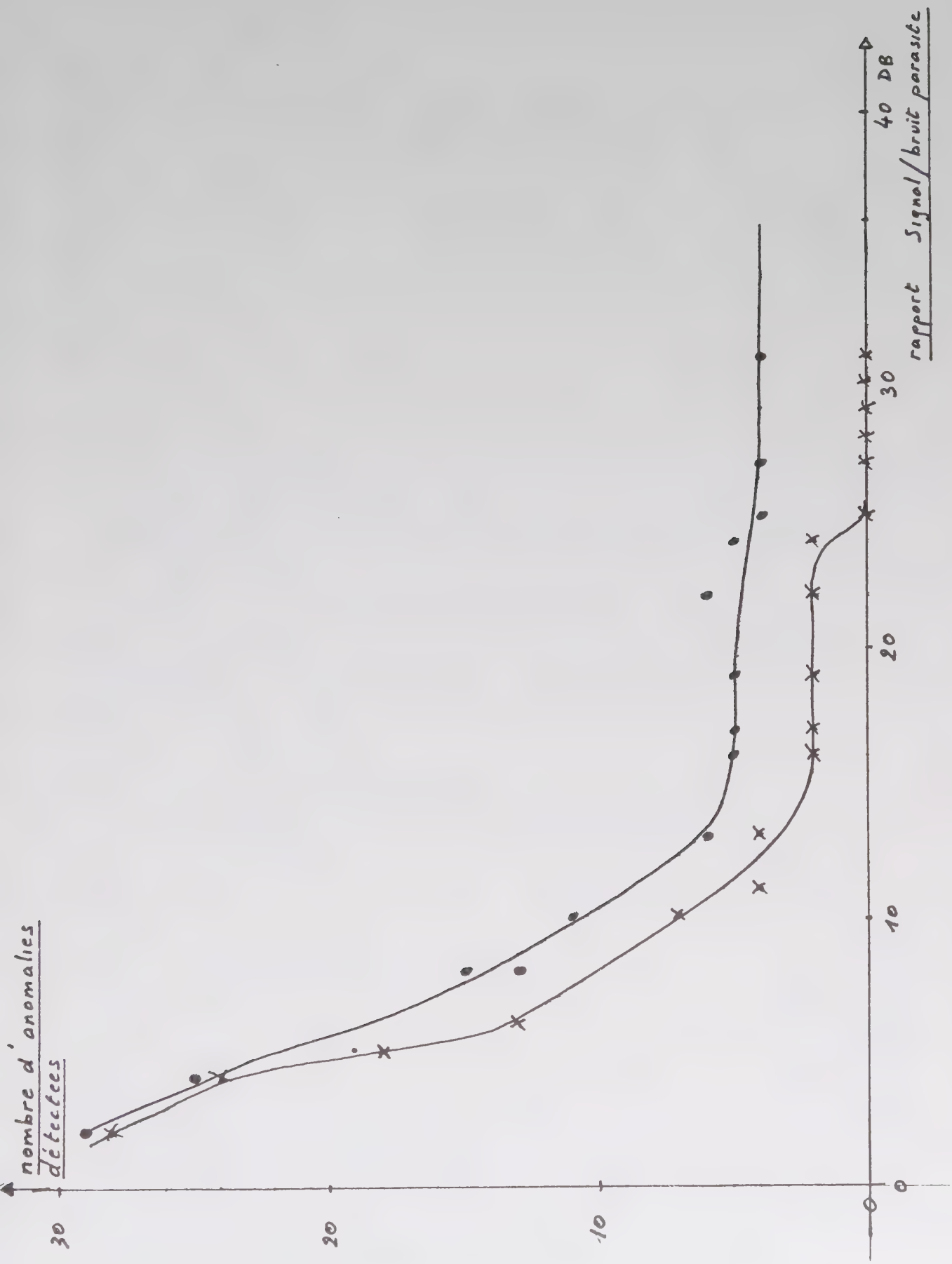


Fig 43: Nombre d'anomalies détectées en fonction du rapport Signal/Bruit blanc parasite (Bruit introduit sur une seule voie)

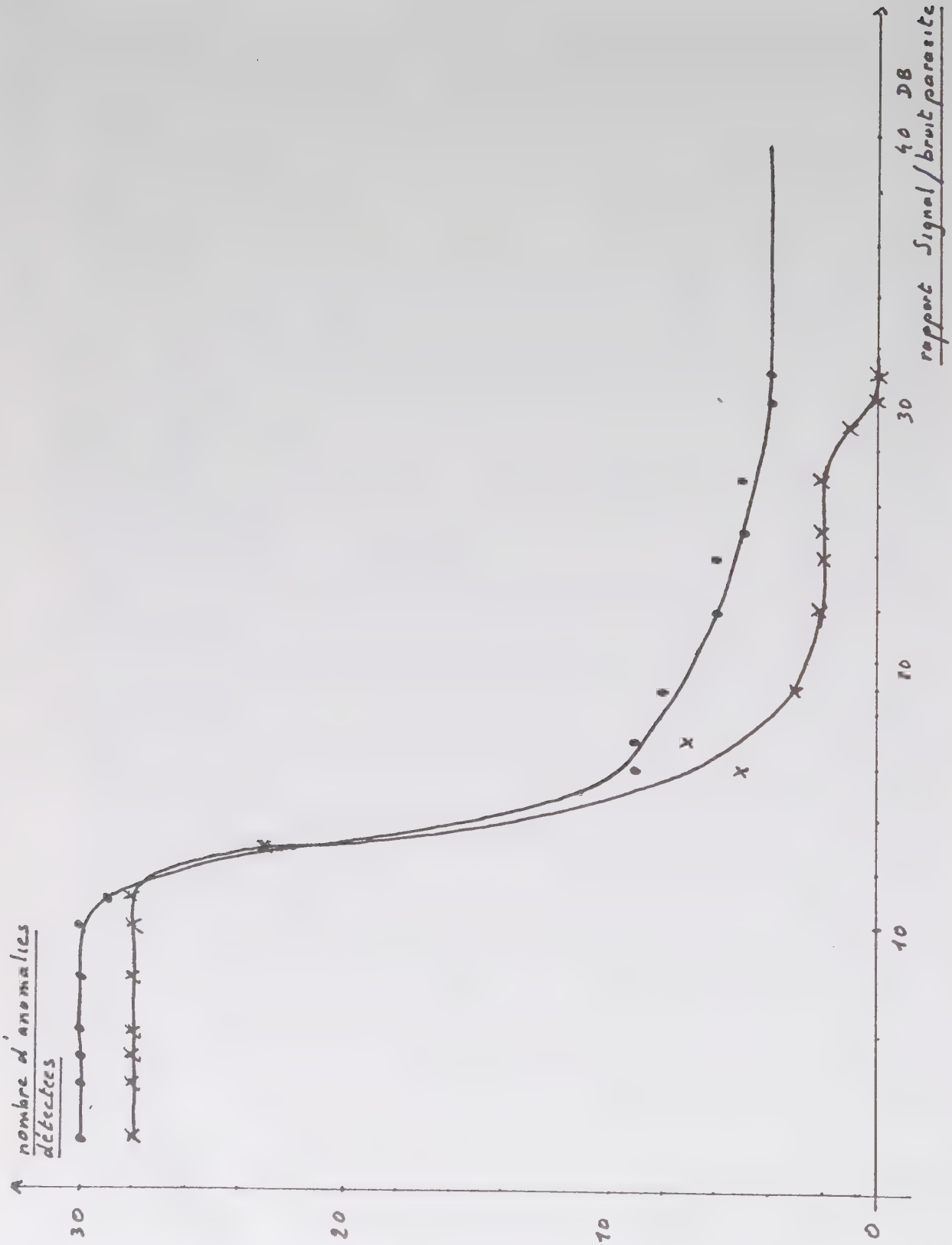


Fig 44: Nombre d'anomalies détectées en fonction du rapport Signal/Bruit blanc parasite
(Bruit introduit sur deux voies)

3.8. - Fonctionnement temps réel

3.8.1. - Apprentissage hors ligne

La durée de l'apprentissage dépend du nombre de formes de la base d'apprentissage, du nombre de groupes trouvés et de la rapidité de convergence des algorithmes itératifs. Ce temps est en moyenne de 3h30* pour une base de 250 formes. Dans le cas d'un apprentissage destiné à une implantation effective nous estimons que la base devrait contenir environ 1000 formes. La durée de l'apprentissage serait alors de l'ordre de 5-6h*. Ces temps qui ne tiennent pas compte des calculs des corrélations peuvent paraître importants mais il faut noter que dans le cas d'une application, l'apprentissage est réalisé une seule fois avant implantation.

3.8.2. - Surveillance en ligne

La simplicité des algorithmes utilisés permet d'envisager une implantation en ligne de la surveillance. Celle-ci pourrait se faire suivant le schéma suivant :

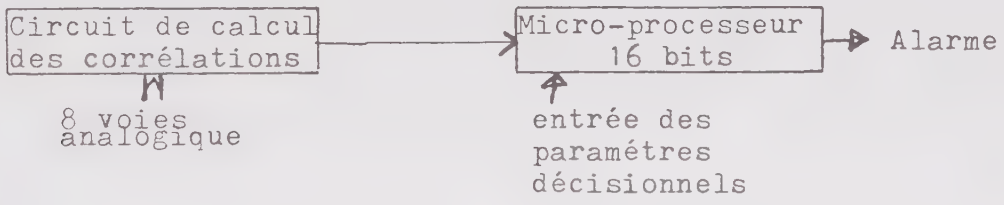


Fig 45: Implantation en ligne

---

* Ces temps correspondent à des implantations sur un mini ordinateur du type SOLAR 16-65. (Calculateur 16 bits). Les fichiers "formes" étant portés par un disque magnétique piloté par le même axe que le disque "système" chaque accès à une forme nécessite un accès disque (50 ms).

A titre d'exemple, sur un microprocesseur du type TMS 990 0 (Texas-Instrument), les temps d'exécution seraient les suivants :

- Calcul des paramètres à partir de la corrélation 10 s.
- Reconnaissance 2 s par groupe et par forme.

Une décision quand au caractère normal ou anormal pourrait donc être prise environ toute les minutes.

4 - ADAPTATIVITE DE LA RECONNAISSANCE DE FORMES

4.1. - Introduction

Le chapitre 2 montre comment un problème de surveillance peut être formulé en terme de reconnaissance de formes sous l'hypothèse de normalité constante. Nous y avons vu que parfois, il est nécessaire d'admettre une certaine évolution des critères de jugement. Le phénomène de veillissement peut imposer une telle évolution. En effet, si l'on est en droit d'attendre d'un processus neuf, un fonctionnement impéccable, il est parfois nécessaire de considérer comme normaux, certains types de fonctionnements imparfaits dans ce même processus ayant vieilli. Notons que ce type d'évolution n'est pas réservé au domaine temporel. Tout autre paramètre peut en être source. Il peut s'agir, par exemple, de la charge d'une machine qui n'a pas le même comportement suivant le travail qui lui est demandé, d'une température, d'une pression atmosphérique... Ces paramètres, rendant nécessaire une indexation des règles de décision peuvent être continus ou discrets, introduire des discontinuités ou pas...

Dans le cas traité au chapitre 3 de la surveillance d'un coeur de réacteur nucléaire une évolution des paramètres en fonction du temps existe. En effet, il existe des différences entre spectres neutroniques de début et spectres neutroniques de fin de cycle. La figure 49 montre quatre spectres calculés dans les mêmes conditions mais à quatre dates différentes du signal émis par une section de détection du réacteur nucléaire de Bugey II. Ces spectres sont calculés sur des fenêtres temporelles de une heure. Une application sur données réelles des algorithmes proposés dans ce chapitre n'a pu toutefois être réalisée. En effet, une telle application suppose un nombre très important d'enregistrements datés. Nous montrerons toutefois au paragraphe 4.4. comment l'algorithme proposé peut être utilisé sur de telles données. L'efficacité de la méthode sera testée sur des données simulées.

4.2. - Approches classiques

Le trait commun des approches qui suivent est d'effectuer un apprentissage sur un ou plusieurs cycles complets. Elles nécessitent donc la connaissance d'un certain nombre d'expériences à chaque instant et supposent que d'un cycle à l'autre l'évolution de la normalité est sensiblement la même. On peut distinguer deux types d'apprentissage et deux types de reconnaissance.

- Apprentissage statique

Le processus est lancé p fois. Lors de chaque cycle i $i = 1, p$ et à chaque instant t_k on mémorise x_k^i ainsi que la décision prise par le professeur $S(x_k^i)$. Une base B d'apprentissage est ainsi formée : $B = \bigcup_{t_k} \{ (x_k^i, S(x_k^i)) \ i = 1,p\}$ Les règles de décisions S_k sont générées à l'aide des sous bases :

$$B_k = \bigcup_{j \in [k-N,k-1]} \{ (x_j^i, S(x_j^i)) \quad i = 1,p\}$$

- Apprentissage dynamique

La détermination des règles de la décision S_k peut parfois être effectuée en ligne de façon récursive :

$$S_k = f\left(S_{k-1} + \varepsilon(\{ (x_{t_k}^i, S(x_{t_k}^i)) i=1,p\}, \{ (x_{t_{k-N}}^i, S(x_{t_{k-N}}^i)) i=1,p\})\right)$$

Terme correcteur

rappelons que S_k est une fonction de $\mathbb{R}^n \to \{a,n\}$

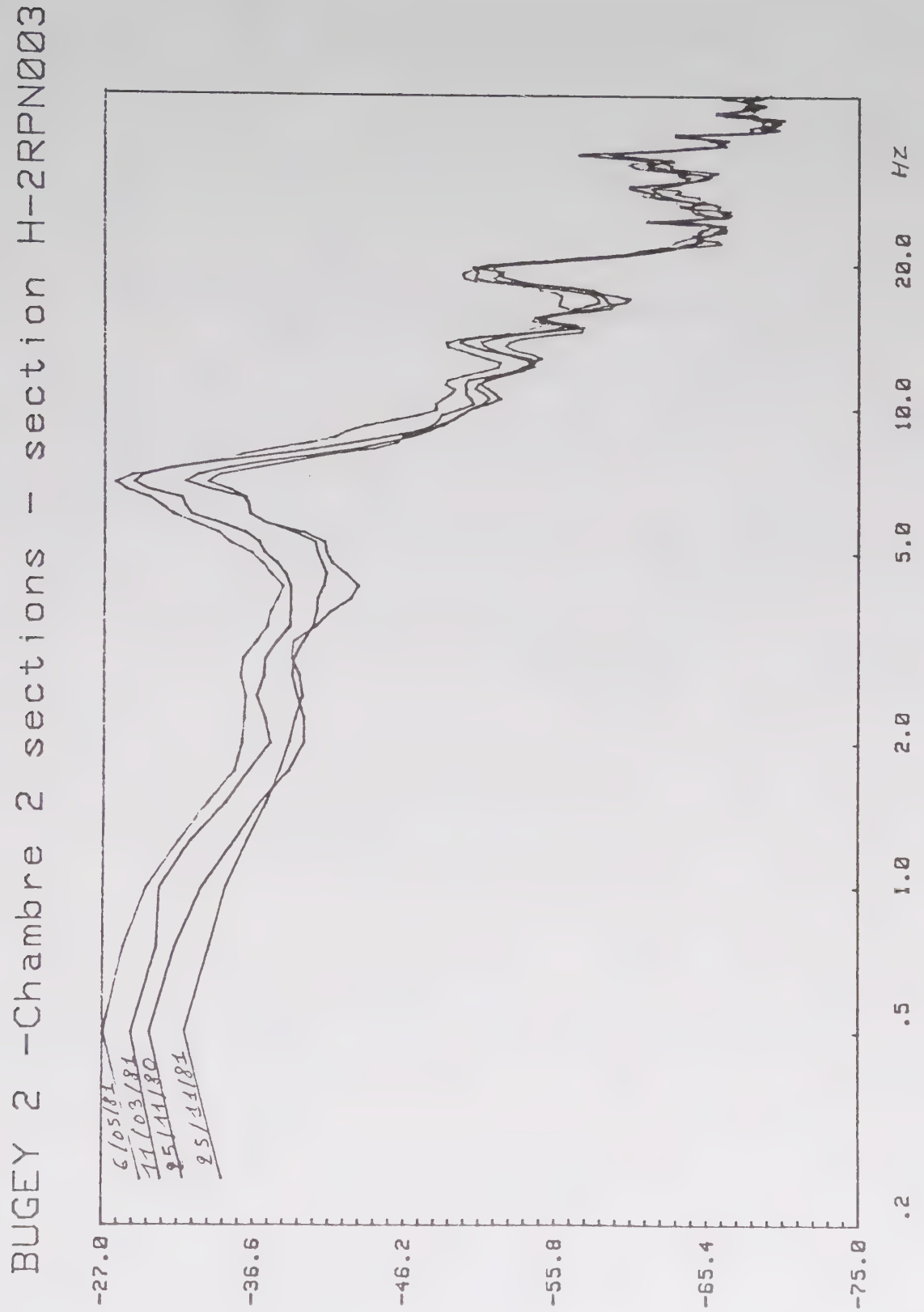


Fig 49: Evolution des spectres neutroniques du coeur de la centrale nucléaire de BUGEY II

- Reconnaissance en boucle ouverte

Les règles de décision S_k calculées lors de la phase d'apprentissage sont définitivement admises. Lors d'un cycle ultérieur la reconnaissance est à chaque instant :

$$x_k \rightarrow S_k(x_k)$$

- Reconnaissance en boucle fermée

Lors d'un cycle ultérieur, au temps t_k, on observe x_k la décision prise est $S_k(x_k)$. Sous réserve qu'un certain critère de fiabilité soit satisfait le couple $(x_k, S_k(x_k))$ est inscrit dans la base B. Avant chaque nouveau cycle une nouvelle phase d'apprentissage est lancée. Les décisions prises lors des cycles précédents sont prises en compte pour déterminer les nouvelles fonctions de décision. Cette nouvelle détermination peut être complète ou n'être qu'une correction apportée aux règles établies.

$$S_k^{(n+1)} = g\left(S_k^{(n)},\ \varepsilon((x_j, S_j^{(n)}(x_j))\ \ j=k-N \rightarrow k)\right)$$

Le schéma d'une procédure de surveillance avec apprentissage statique et décision en boucle ouverte est le suivant :

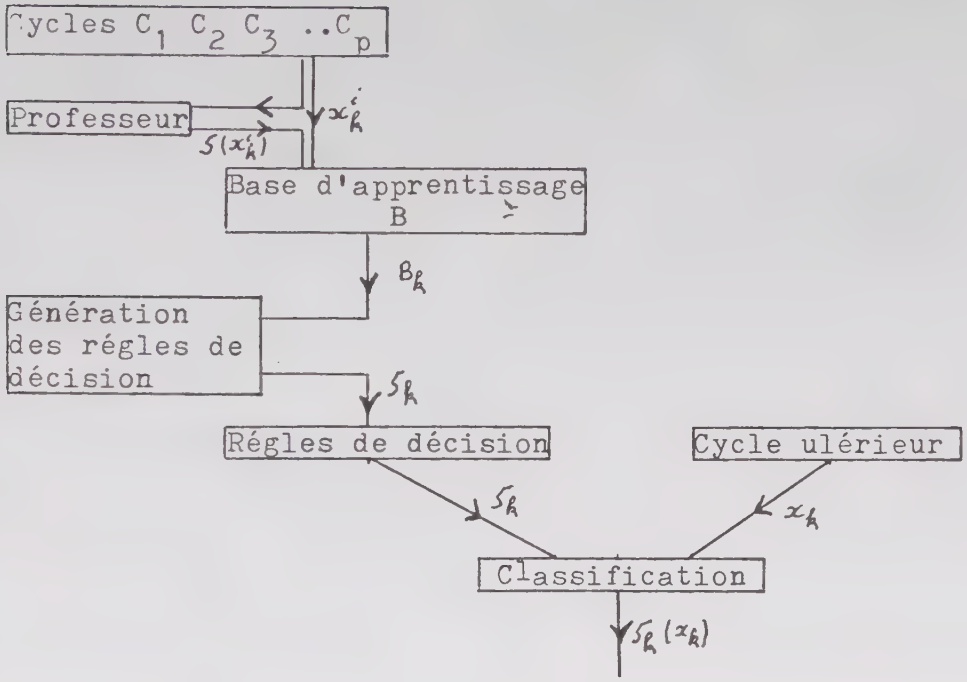


Fig 50: Apprentissage statique, Reconnaissance en boucle ouverte

La génération des règles de décisions s'effectue une fois pour toute. L'apprentissage et la reconnaissance sont complètement découplés.

Le schéma d'une procédure de surveillance avec apprentissage statique et décision en boucle fermée est le suivant :

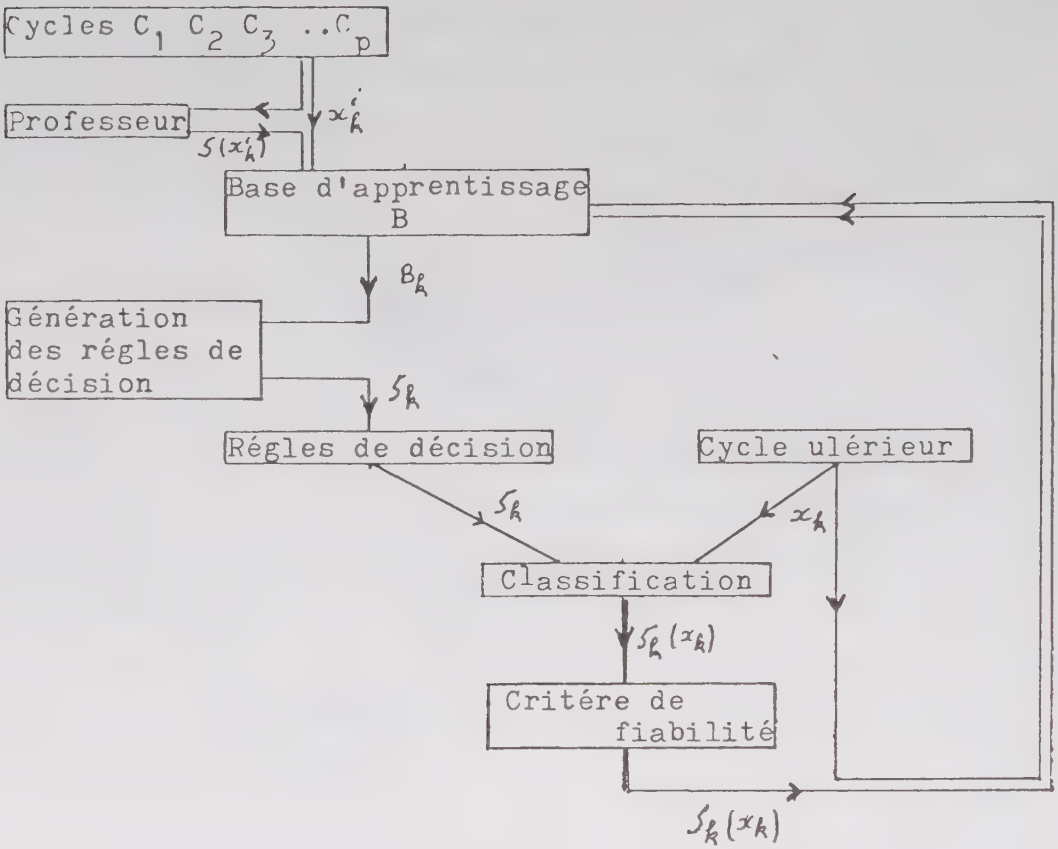


Fig 51: Apprentissage statique
Décision en boucle fermée

La génération des règles de décision s'effectue une fois avant chaque nouveau cycle.

Les schémas de procédures de surveillance avec apprentissage dynamique seraient semblables. Dans le cas d'un apprentissage dynamique avec décision en boucle fermée la procédure serait la suivante : à chaque instant t_i une règle de décision S_i est générée. Elle sert à classer l'observation x_i. Lorsque cette classification est fiable le couple $(x_i, S_i(x_i))$ est introduit dans la base B. Cette nouvelle donnée va donc être considérée lors des déterminations des règles S_{i+1} ... S_{i+N}. L'importance relative des données issues du cycle en cours relativement à celles provenant des cycles précédents peut être pondérée.

4.3. - Reconnaissance adaptative

Les méthodes précédentes supposent que la notion de normalité évolue de la même façon lors de cycles différents. Cette supposition n'est pas toujours vérifiée. D'autre part, la mémorisation à chaque instant des décisions d'apprentissage et des observations sur p cycles est souvent infaisable. Ces considérations nous ont fait imaginer la procédure de surveillance de la figure 52 dont la particularité est de ne demander qu'un apprentissage (professé ou non) en début de cycle. Ce schéma comprend deux phases. La première est un apprentissage classique sur une base de données composée d'épreuves de début et éventuellement de plusieurs cycles. En effet, s'il est douteux que l'évolution des critères se fasse de la même façon d'un cycle à l'autre, il est souvent admissible de les considérer identiques en début de chaque cycle. Cet apprentissage génére la première règle de décision. La deuxième phase est une reconnaissance en boucle fermée : au temps t_i, la forme x_i est reconnue grâce à la règle de décision S_i, si cette reconnaissance satisfait un certain critère de fiabilité (propre à la méthode), le couple $(x_i, S_i(x_i))$ est introduit dans la base de données et une nouvelle règle de décision est générée.

4.4. - Reconnaissance adaptative par modélisation sphérique

4.4.1. - Présentation de la méthode

La méthode que nous proposons ci-dessous suppose que chaque forme est représentée par un vecteur de $\mathbb{R}^n$ et que l'on dispose dans cet espace d'une norme mesurant les dissemblances entre formes.

Son principe est conforme à celui décrit au §.4.3. Il comprend deux phases :

La première est un apprentissage avec professeur qui détermine pour chaque classe de formes un ensemble de sphéres dont la réunion forme un modèle géométrique de l'ensemble des points de $\mathbb{R}^n$ représentatifs de la classe.

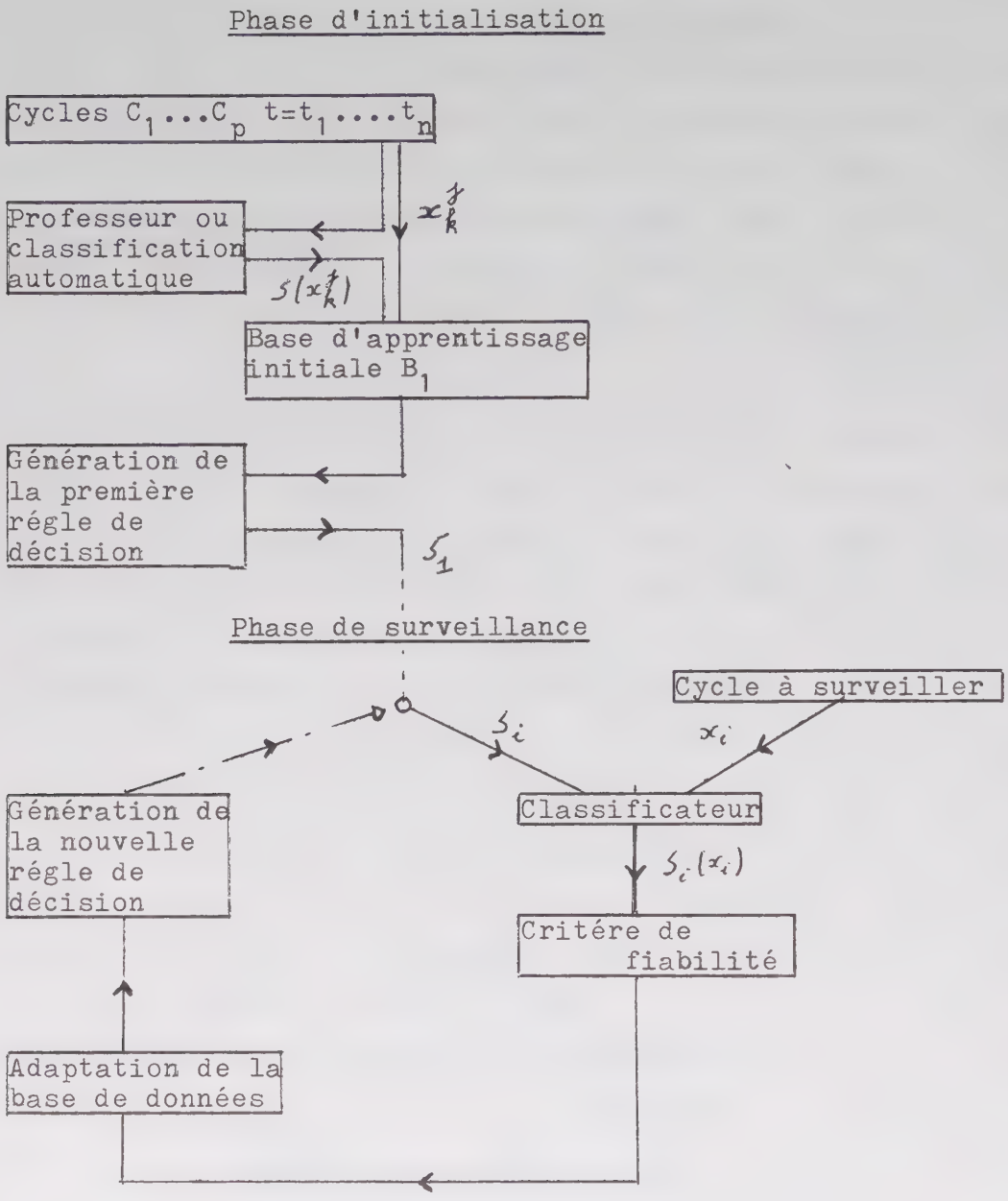


Fig 52: Reconnaissance adaptative

La deuxième est une reconnaissance adaptative qui classe chaque nouvelle forme inconnue sur la base des modèles pré-établis puis adapte, à l'aide de cette décision, les modèles en modifiant les centres, les rayons et le nombre des sphéres les composant.

4.4.2. - La phase d'apprentissage

Etant donné un ensemble ω de N points de $\mathbb{R}^n$ partionné en classes ω_i $i = 1,M$ et une norme que nous noterons $\| \ \|$. On cherche à établir une modélisation de chaque classe ω_i en un certain nombre de sphéres vérifiant les deux propriétés minimales suivantes :

- L'ensemble des sphéres modélisant une classe doit contenir plus d'un certain pourcentage de ses points (le pourcentage minimum sera noté "Pour").

- Dans chaque sphére modélisant une classe, le rapport du nombre de formes appartenant à cette classe sur le nombre de formes ne lui appartenant pas doit être supérieur à un rapport minimal "R_v" que nous appellerons par abus rapport de vraisemblance.

L'intérêt d'une telle modélisation est de permettre l'établissement de règles de décision simples à mettre en oeuvre. En effet, soit une forme inconnue à classer. Le calcul des distances de cette forme aux centres des sphéres et la comparaison de ces distances aux rayons permettent de déterminer les sphéres auxquelles elle appartient et par la même la classe.

L'organigramme de la figure 53 montre pour une classe ω_o, le principe de fonctionnement de cette phase d'apprentissage. La procédure est identique pour chacune des autres classes. Remarquons que la modélisation d'une classe est effectuée par opposition aux autres. Le nombre de classes en présence doit donc être au moins égal à 2. Dans le cas d'une seule classe l'algorithme devient trivial et modélise cette classe par la plus petite sphére centrée au centre de masse et contenant tous les points.

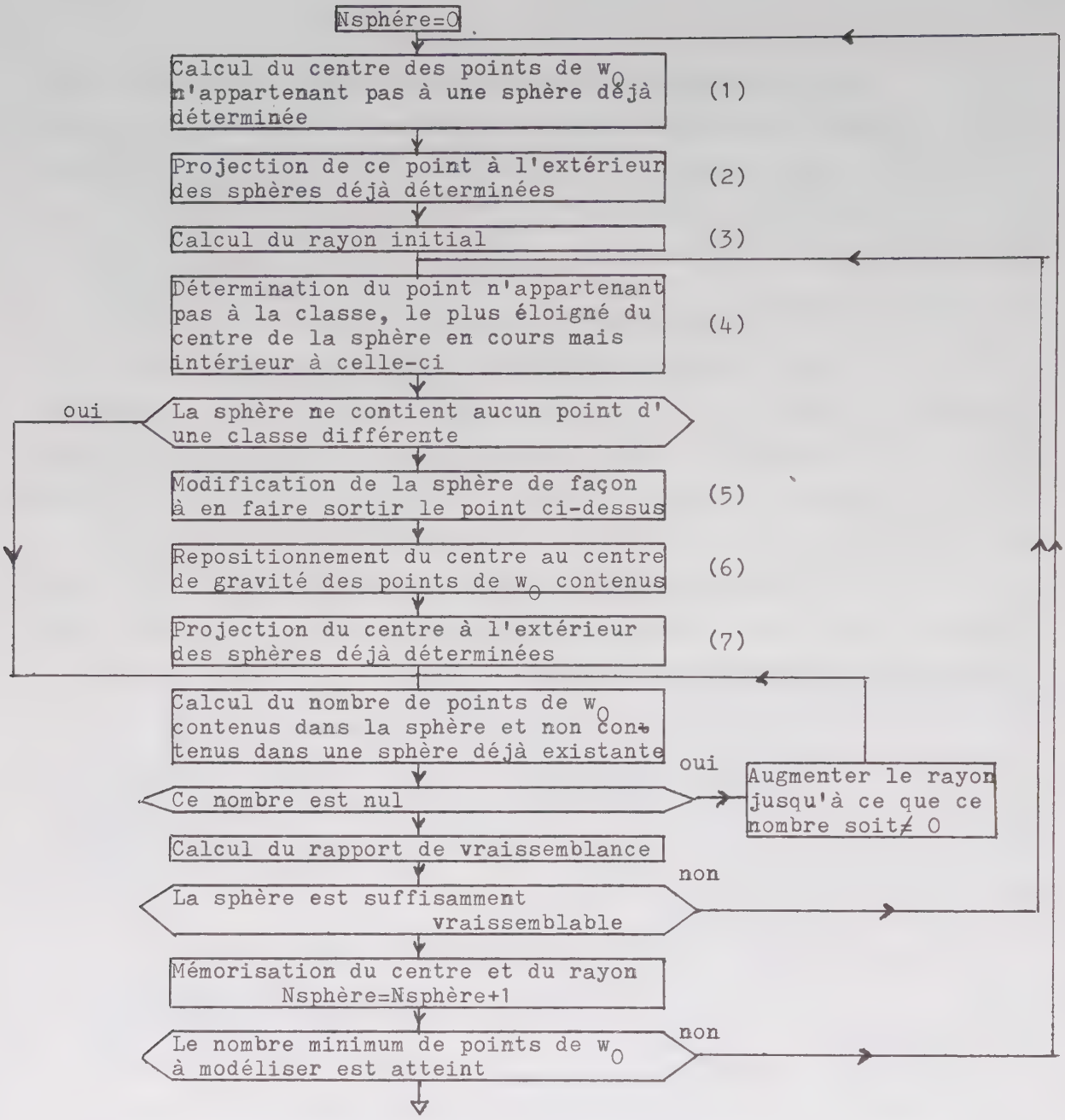


Fig 53: Modélisation d'une classe

La projection de l'étape 2 se fait de façon radiale et centrifuge en considérant les sphéres dans l'ordre de leur indice et de façon itérative jusqu'à ce que le point transformé soit extérieur à toutes les sphères. Les cas d'indétermination sont liés par une perturbation de la position du point initial.

L'étape 3 initialise le rayon par la plus grande distance séparant le centre en cours d'un point ω_o non modélisé. La sphére en cours de détermination contient alors tous les points de ω_o non encore modélisés.

La déformation minimale de la sphère de l'étape $\overline{S}$ est effectuée de la façon suivante : soit M le point à faire sortir :

$$\overrightarrow{OC} \longrightarrow \overrightarrow{OC} + \frac{1}{2} \frac{\overrightarrow{MC}}{||MC||} (R - ||MC||)$$

$$R \longrightarrow (R + ||MC||)/2$$

Note : On dira qu'un point appartient en propre à une sphère si ce point lui appartient et n'appartient à aucune sphère déterminée précédemment et modélisant la même classe.

4.4.3. - Résultats expérimentaux en simulation

Cet algorithme a fait l'objet d'une implantation informatique dont l'organigramme est donné figure 54. Afin de pouvoir illustrer son fonctionnement nous l'avons fait opérer sur données artificielles en dimension 2. Les figures 55 à 61 montrent des modélisations obtenues lorsque les classes ont des supports disjoints. Les sphéres y sont numérotées suivant leur ordre d'apparition. On remarque qu'en général les zones de forte densité sont modélisées en priorité. Ce résultat était escompté et prévisible : en effet, sauf cas exceptionnel, les étapes 4.5. et 7 n'apportent que de légères perturbations à la position de la sphère en cours. Ces perturbations sont de nature quasi-aléatoire. Par contre, l'étape 6 a une tendance prouvée à déplacer le centre vers la zone de plus forte densité.

Les hachures horizontales représentent les zones de décision pour la classe 1 et les hachures verticales celles de la classe 2. Ces exemples montrent l'intérêt de cette méthode de modélisation qui permet, dans des cas non traitables par les algorithmes classiques de discrimination linéaire (voir annexe bibliographique) d'assurer une reconnaissance parfaite lorsque les supports des classes sont disjoints.

Les figures 63 et 65 montrent les modélisations obtenues sur les données des figures 62 et 64. Ces données proviennent de deux distributions gaussiennes dont les centres sont séparées d'une distance unitaire et d'écart type radial 0.3. Ces modèles ont été obtenus pour un pourcentage de recouvrement de 85% et un rapport de vraisemblance de 5. Les règles de décision sont représentées par un trait pointillé pour $x \in \omega_1$ et par un trait continu pour $x \in \omega_2$. Dans la zone comprise entre ces deux traits il y a rejet. La décision baysienne théorique qui dans notre cas est la médiatrice au segment joignant les deux centres est marquée par la lettre B. On constate qu'elle passe dans la zone de rejet.

4.4.4. - La phase de reconnaissance adaptative

L'organigramme de la figure 66 montre le principe de cette reconnaissance.

L'étape 1 supprime les points (formes) situés dans une sphère entrant dans la modélisation d'une classe à laquelle ils n'appartiennent pas. A la fin de cette étape tous les points d'apprentissage situés dans une sphère entrant dans le modèle d'une classe ω_i appartiennent à cette classe.

Cette étape ne fait pas réellement partie de la phase de reconnaissance.

L'étape 2 considère une forme inconnue à classer et détermine les sphères auxquelles elle appartient.

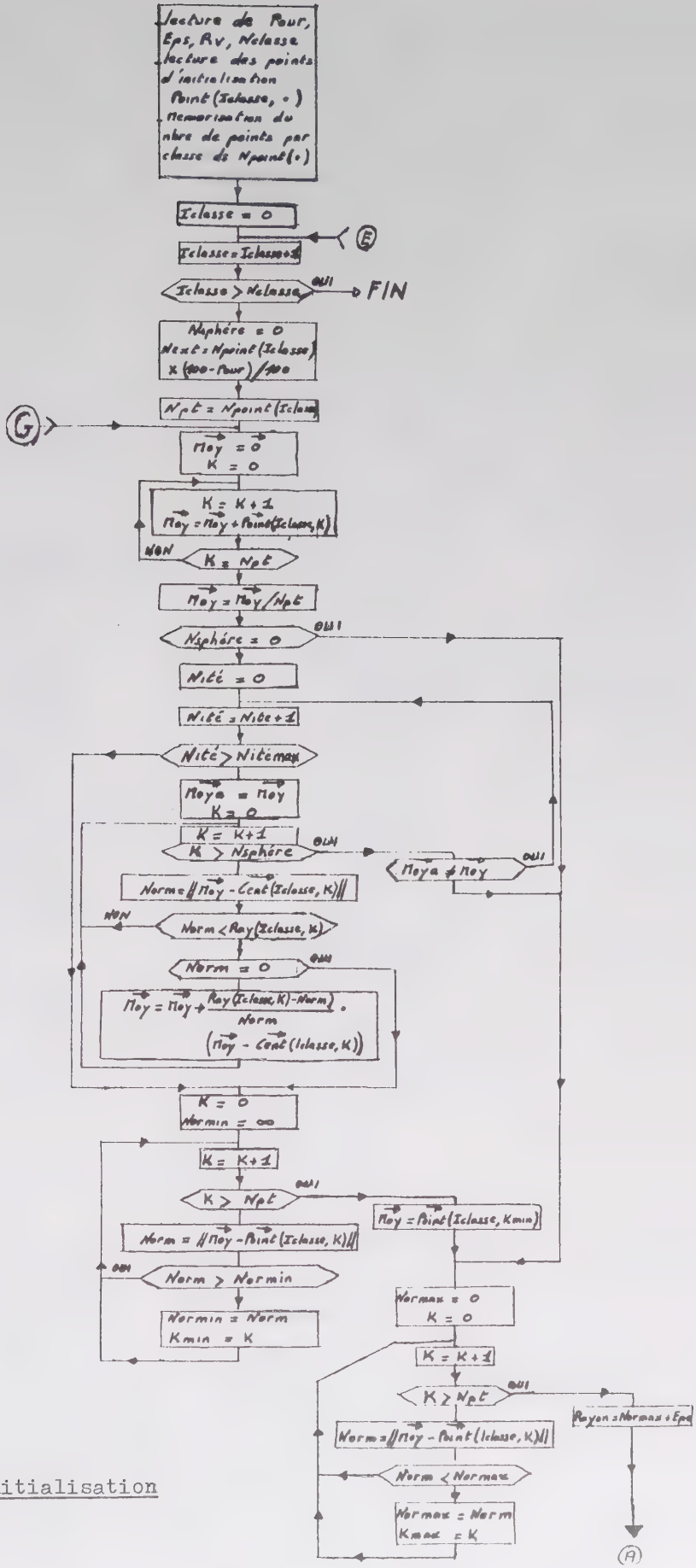


Fig 54: Initialisation

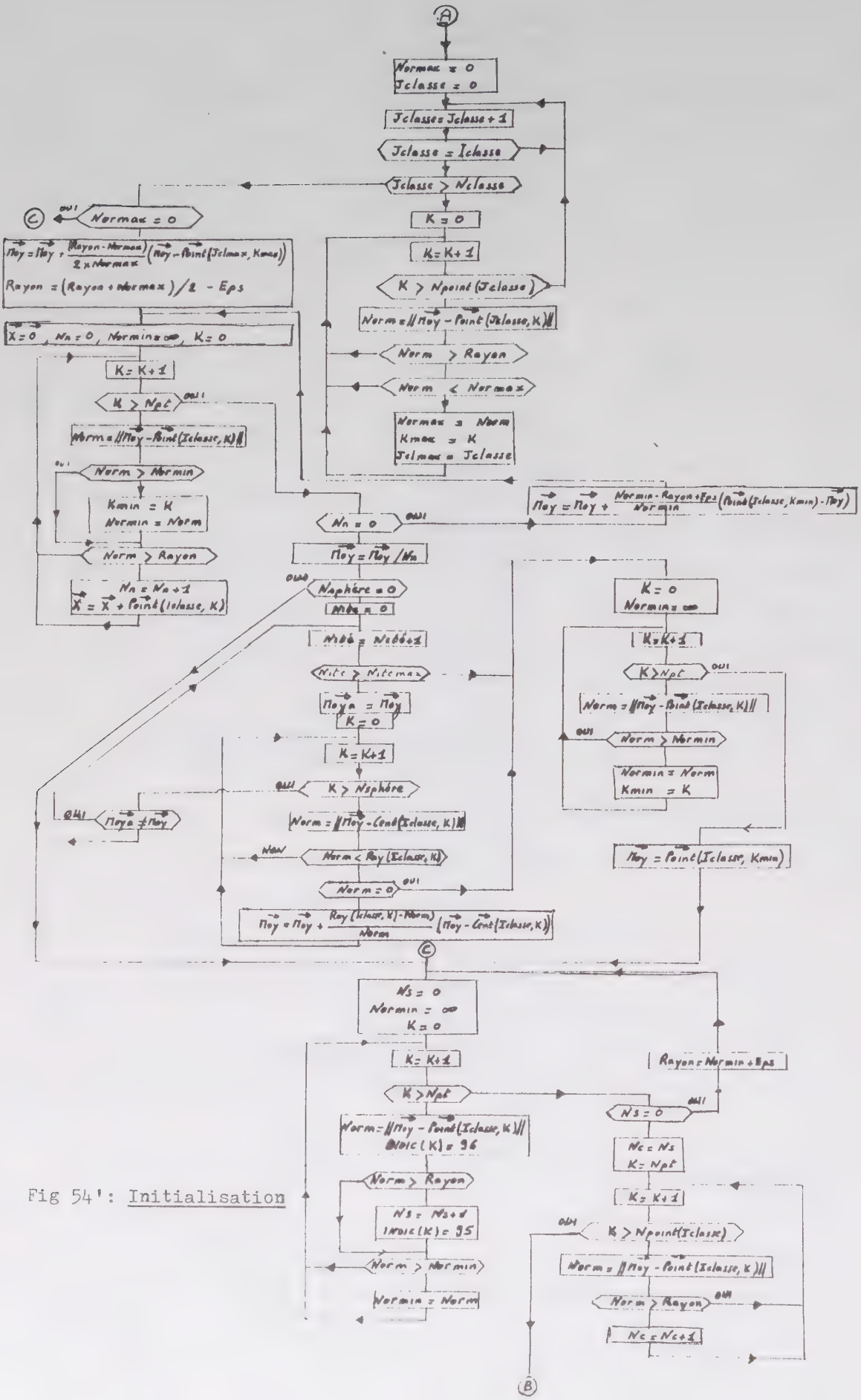
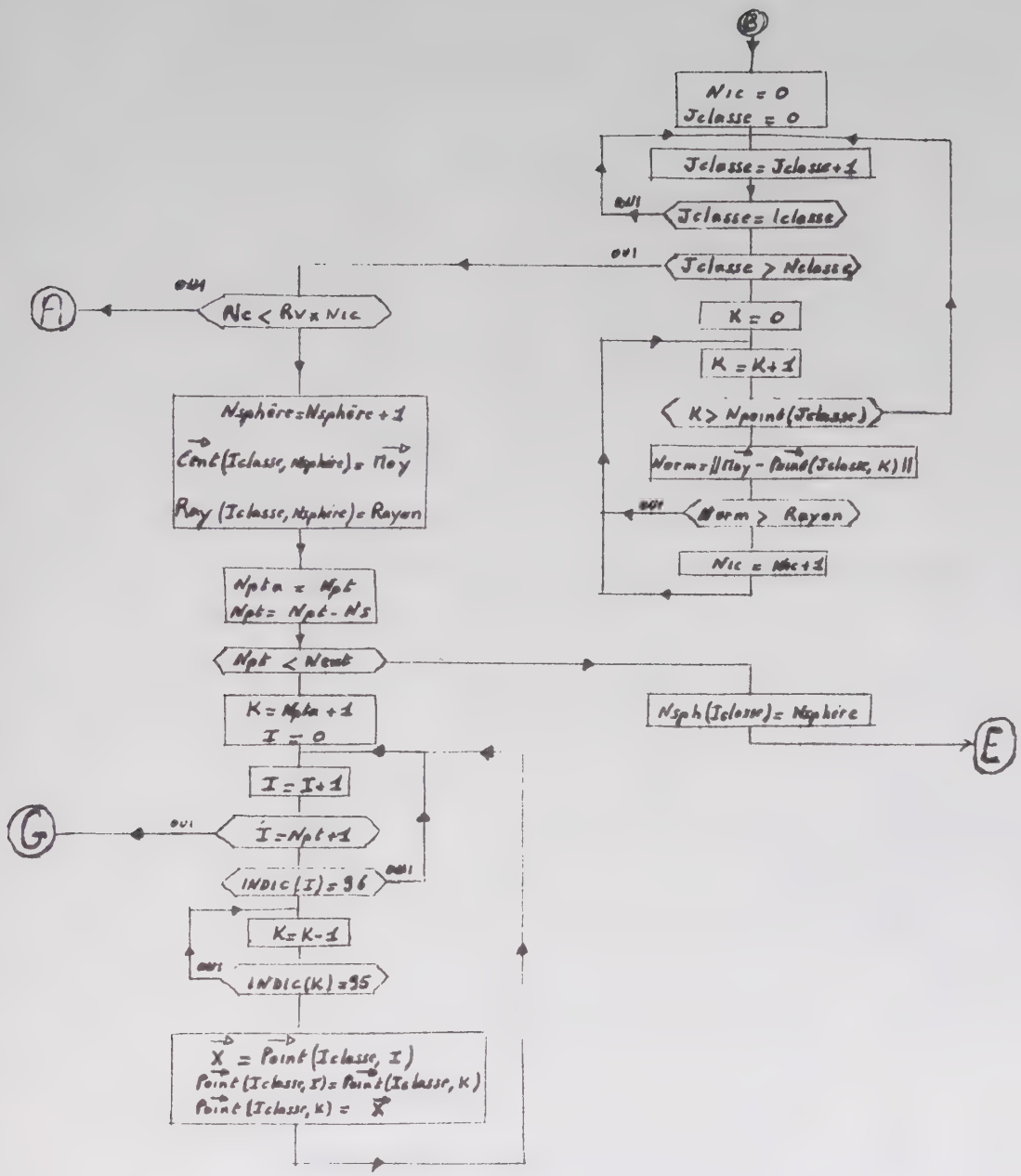


Fig 54': Initialisation



Liste des variables utilisées

Pour	: Pourcentage de recouvrement	Norm	: Distance entre deux points
Rv	: Rapport de vraissemblance	Rayon	: Rayon en cours
Nclasse	: Nombre de classes		
Point (J, I)	: Coordonnees du J^{ième} point de la classe I		
Npoint (I)	: Nombre de pts classe I		
Iclasse	: Indice de classe		
Nsphère	: Nombre courrant de sphères		
Next	: Nombre de pts pouvant ne pas être modélisés		
Npt	: Nombre de pts restant à modeliser		
Moy	: Centre en cours		
K	: Indice de spheres ou de points		
Nité	: Nombre d'iterations		
Nitemax	: " " permises		
Moya	: Memorisation de Moy		

Fig 54'': Initialisation

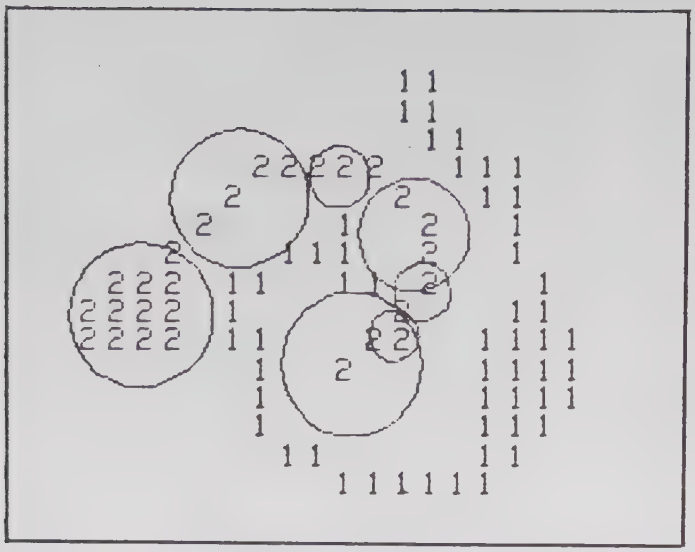


Fig 55

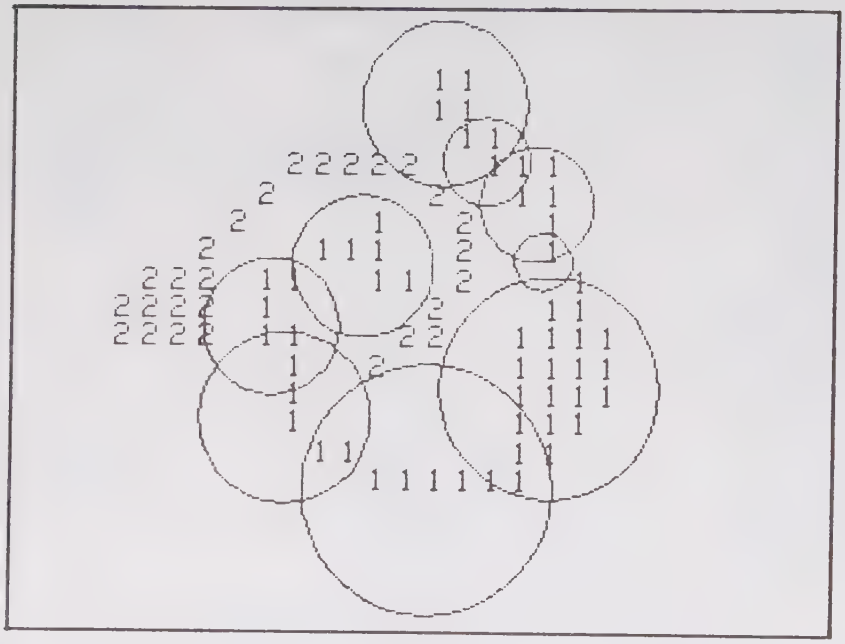


Fig 56

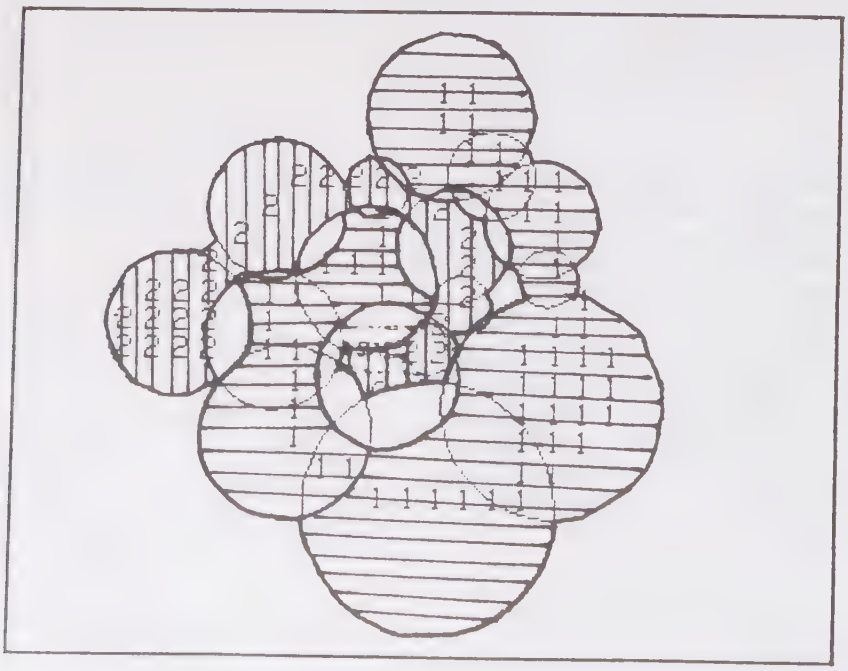


Fig 57

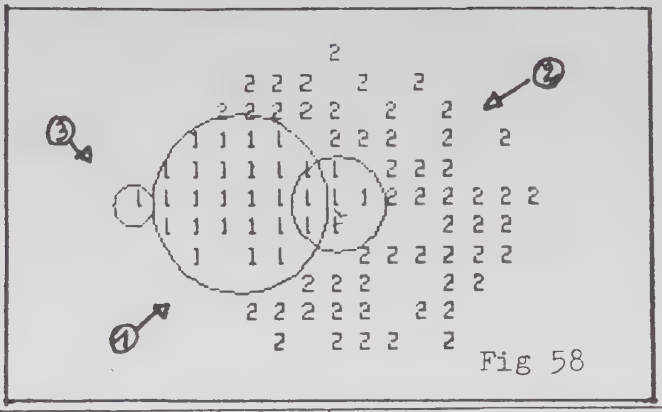


Fig 58

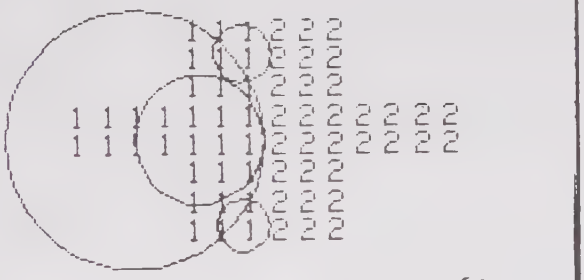


Fig 60

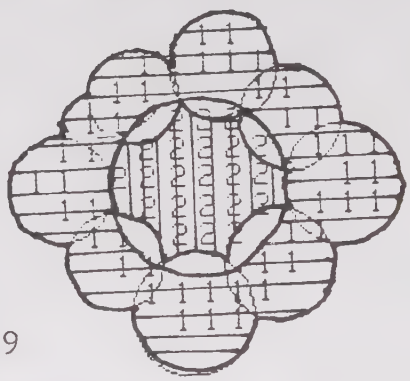


Fig 59

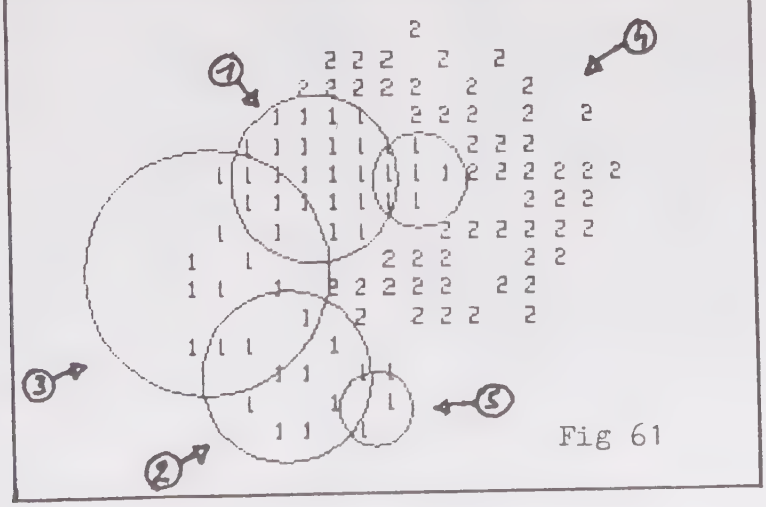


Fig 61

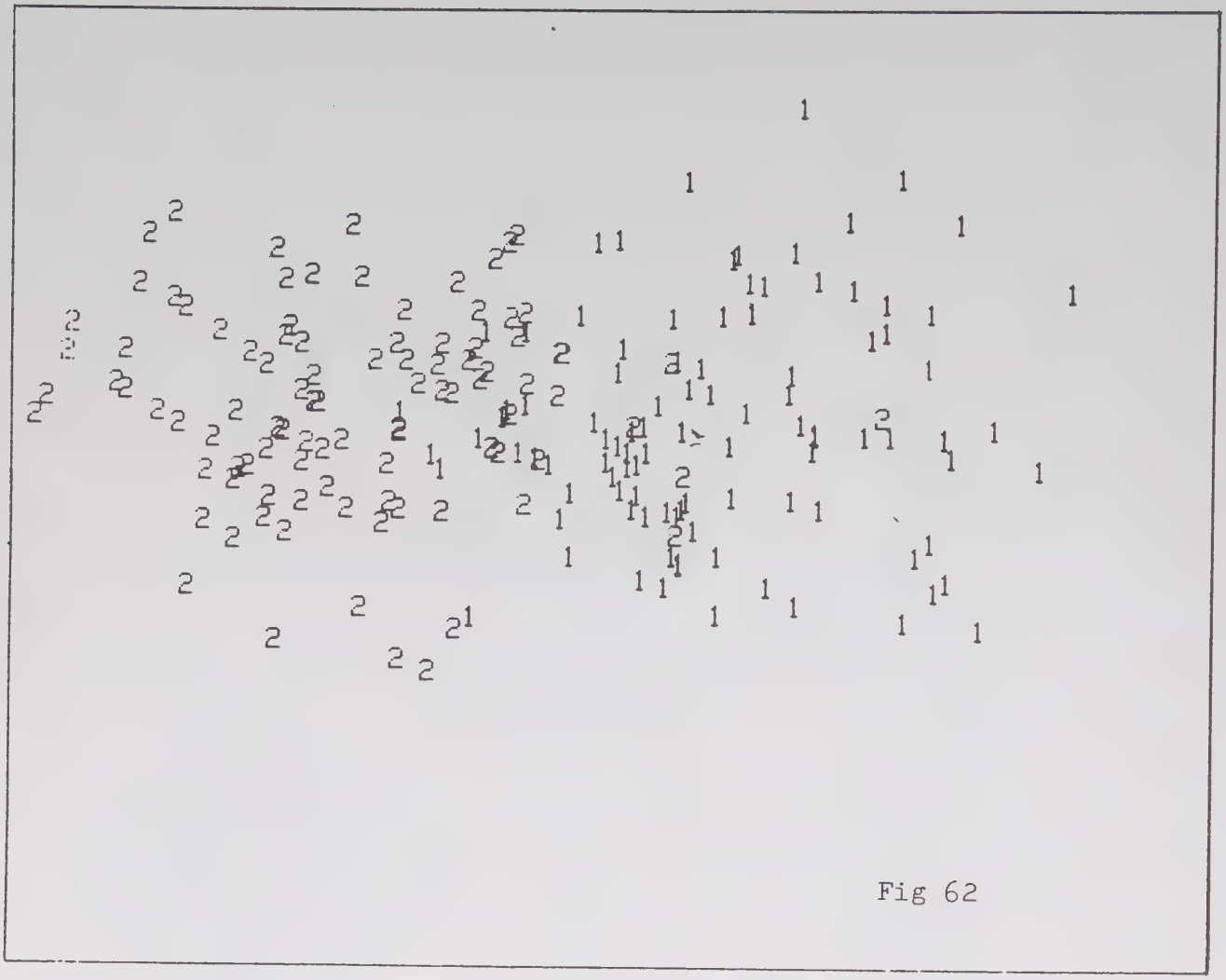


Fig 62

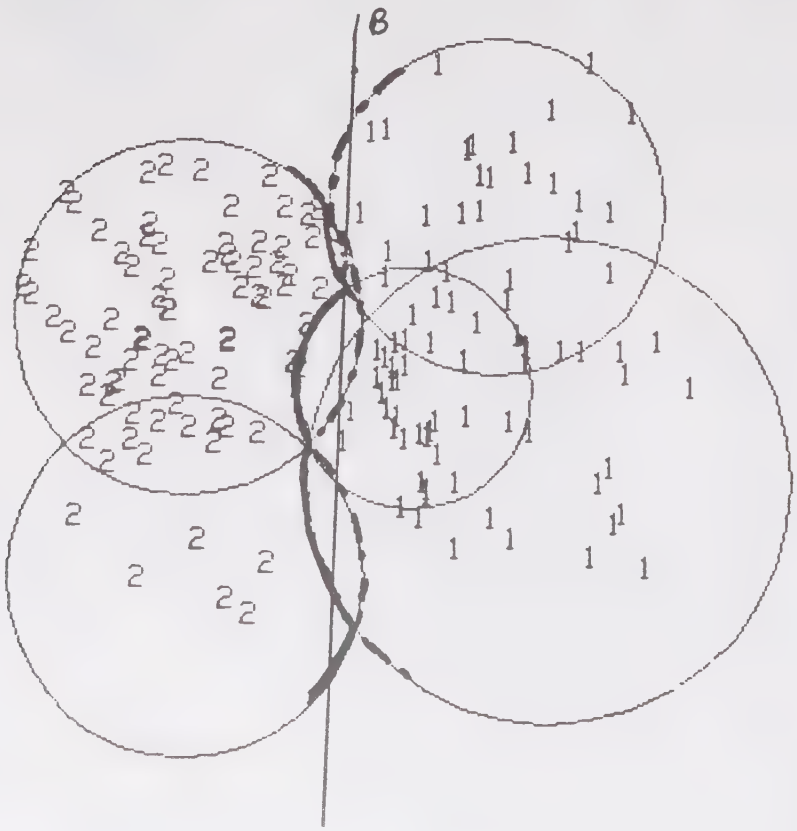


Fig 63

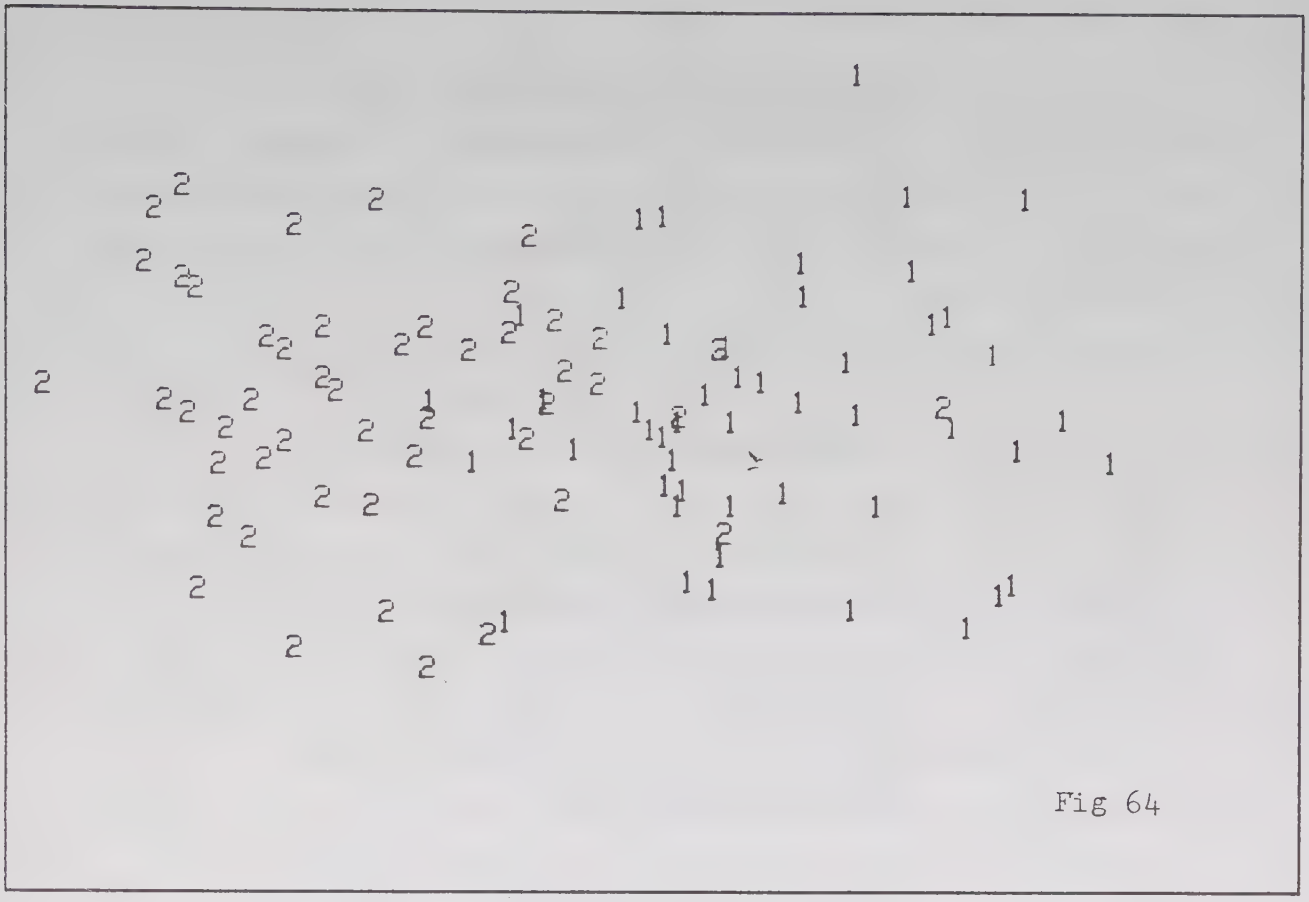


Fig 64

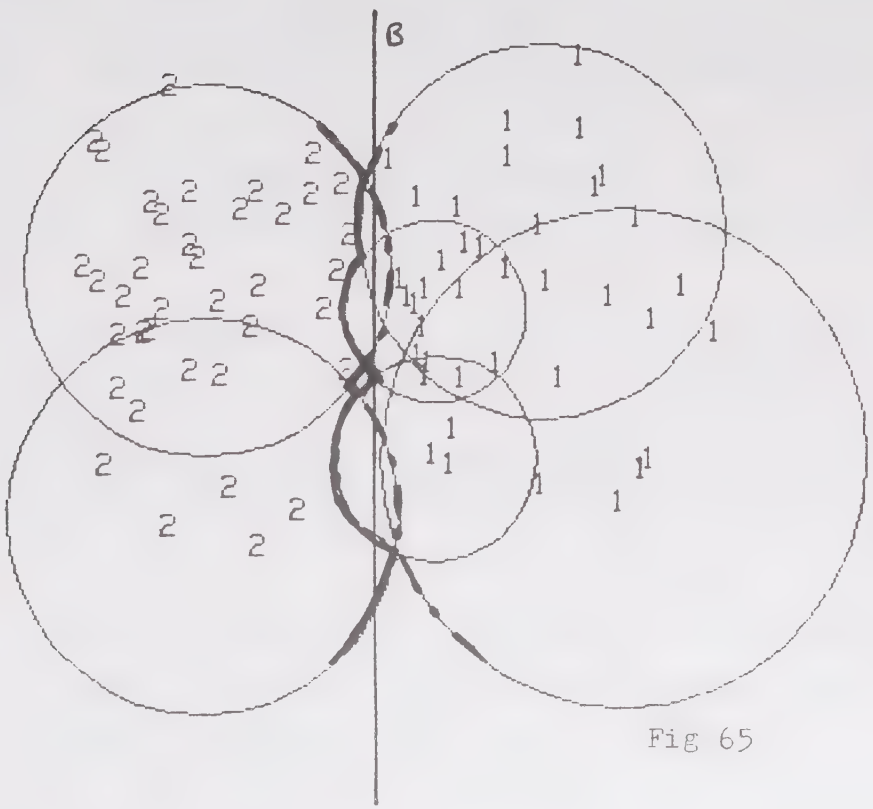


Fig 65

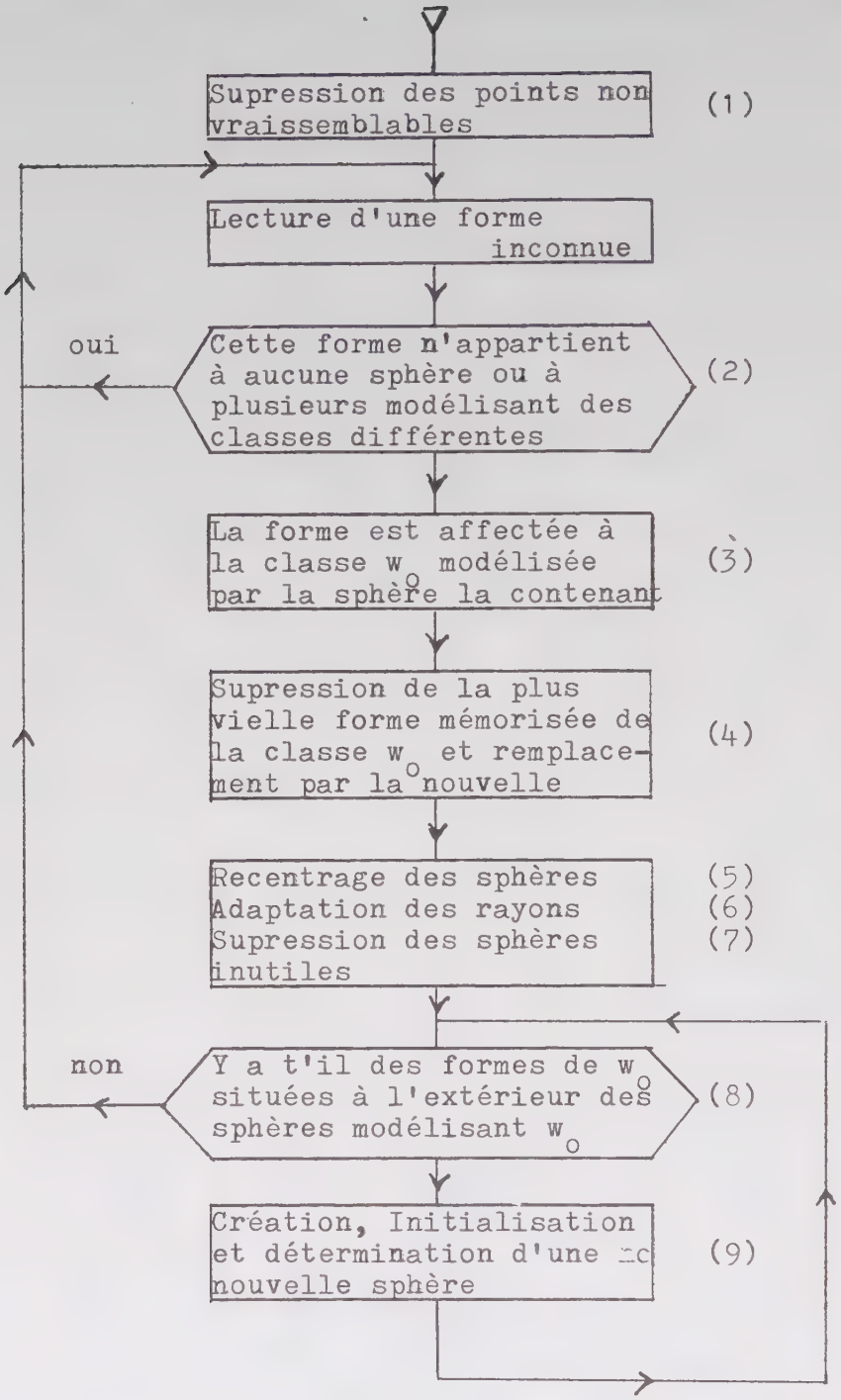


Fig 66: Adaptation des modéles et reconnaissance

Si elle appartient à des sphères modélisant des classes différentes ou si elle n'appartient à aucune sphère l'algorithme ne la considère pas comme suffisamment vraisemblable pour adapter ses paramètres. Une décision au plus proche voisin ou sur un critère de distance au plus proche centre est toutefois possible. Cette décision en seconde main n'aura pas d'influence sur la suite.

Les étapes 3 et 4 déterminent la classe d'appartenance (ω_o) et mémorisent la forme. La plus vieille forme mémorisée dans la même classe est supprimée de façon à ce que le nombre de points mémorisés par classe soit constant. Ce nombre dépend de la taille de la base d'apprentissage et de l'étape ci-dessus.

L'étape 5 recentre chaque sphère modélisant ω_o au centre des points lui appartenant en propre (il y a eu un point de supprimé et un point d'ajouté par l'étape 4).

L'étape 6 adapte les rayons de chacune des sphères ci-dessus comme suit : pour chaque sphère on calcule la plus grande distance séparant son centre d'un point lui appartenant en propre. Soit Rint cette distance. On calcule également la plus petite distance séparant le centre d'un point n'appartenant pas à ω_o. Soit Rext cette distance. Si Rext est inférieur à Rint le rayon est posé égal à Rext. Dans le cas contraire le rayon est égalé à la demi-somme de ces deux grandeurs.

L'étape 7 détermine pour chaque sphère modélisant ω_o le nombre de points lui appartenant en propre c'est à dire le nombre de points de ω_o qu'elle contient et qui ne sont pas contenus dans une sphère d'indice inférieur (ayant été initialisée avant elle). Si ce nombre est nul la sphère est supprimée.

L'étape 8 vérifie que, malgré les modifications des étapes 5,6 et 7, tous les points de ω_o appartiennent au moins à une sphère la modélisant.

L'étape 9 est semblable à la phase d'initialisation de la figure 53 à quelques détails près :

- Le nombre de sphères n'est pas réinitialisé à O.
- Le bloc 3 n'initialise pas le rayon de la même façon.

Soit Rint la plus grande distance séparant le centre d'un point de ω_o non affecté à une sphère.

Soit Rext la plus petite distance séparant le centre d'un point n'appartenant pas à ω_o.

Si Rext est supérieur à Rint le rayon est initialisé à Rext dans le cas contraire le rayon est initialisé à la demi somme de ces deux valeurs.

Le critère d'arrêt d'évolution d'une sphère en cours de détermination est de ne contenir que des points ω_o.

4.4.5. - Résultats expérimentaux en simulation

La figure 67 montre l'organigramme détaillé de cette phase de reconnaissance adaptative. Afin de mettre en évidence l'adaptativité des règles de décisions nous présentons ci-dessous une expérience en dimension 2.

Les données présentées à la phase d'apprentissage sont celles de la figure 68. Elles correspondent à 10 épreuves des lois de probabilité suivante : $\sigma = 0,2$ -

$\rho(x/\omega_1) = N(x_1, C)$ $\quad x_1 = (0,0) \quad C = \begin{pmatrix} \sigma^2 & \\ & \sigma^2 \end{pmatrix}$

$\rho(x/\omega_2) = N(x_2, C)$ $\quad x_2 = (-1,0) \quad C = \begin{pmatrix} \sigma^2 & \\ & \sigma^2 \end{pmatrix}$

Les données présentées à la phase de reconnaissance adaptative sont celles de la figure 69. Elles correspondent à 100 épreuves successives des lois de probabilité suivantes :

$$\rho(x/\omega_1) = N(x_1, C)$$

$$\rho(x/\omega_2) = \frac{1}{2} N(x_3(t), C) + \frac{1}{2} N(x_4(t), C)$$

$$x_3(t) = (-\cos \omega t, \sin \omega t)$$

$$x_4(t) = (-\cos \omega t, -\sin \omega t)$$

Le paramètre ω étant calculé pour qu'à la $100^{ème}$ épreuve $x_3(t)$ et $x_4(t)$ soient en (0,1).

La figure 70-1 montre les modèles générés par la phase d'apprentissage. La figure 70-2 marque le moment de la création d'une nouvelle sphère pour modéliser la classe 2 qui possède maintenant deux modes.

Les figures 70-3 à 70-14 montrent l'évolution au cours du temps des modèles. La figure 70-15 montre le moment de la supression d'une sphère. La figure 70-16 est la modélisation finale. Au cours de cette expérience et malgrè l'adaptation accomplie, il n'y a eu aucune erreur de classification. Il y a eu deux rejets qui ont été correctement classés en 2ème ressort par la méthode du plus proche voisin.

La même expérience a été reconduite pour des valeurs différentes de l'écart type σ. Jusqu'à la valeur $\sigma = 0,3$ correspondant aux données de la figure 71, l'algorithme converge sans précaution particulière. Pour des valeurs plus grandes de l'écart type une optimisation du nombre de points d'apprentissage relativement à la vitesse d'évolution de la position des modes de la classe 2 est nécessaire : ce nombre doit être suffisamment petit pour que les modèles puissent suivre l'évolution et suffisamment grand pour qu'une erreur de classification ne remette pas en cause leur validité. Nous avons obtenu convergence jusqu'à $\sigma = 0.5$. Le nombre d'erreurs commises est alors 20% supérieur au nombre d'erreurs qui auraient été commises par un classificateur baysien théorique optimal à chaque instant. L'évolution des modèles est plus complexe que celle montrée figure 70 : le nombre de sphères modélisant la classe 2 oscillant entre 5 et 6.

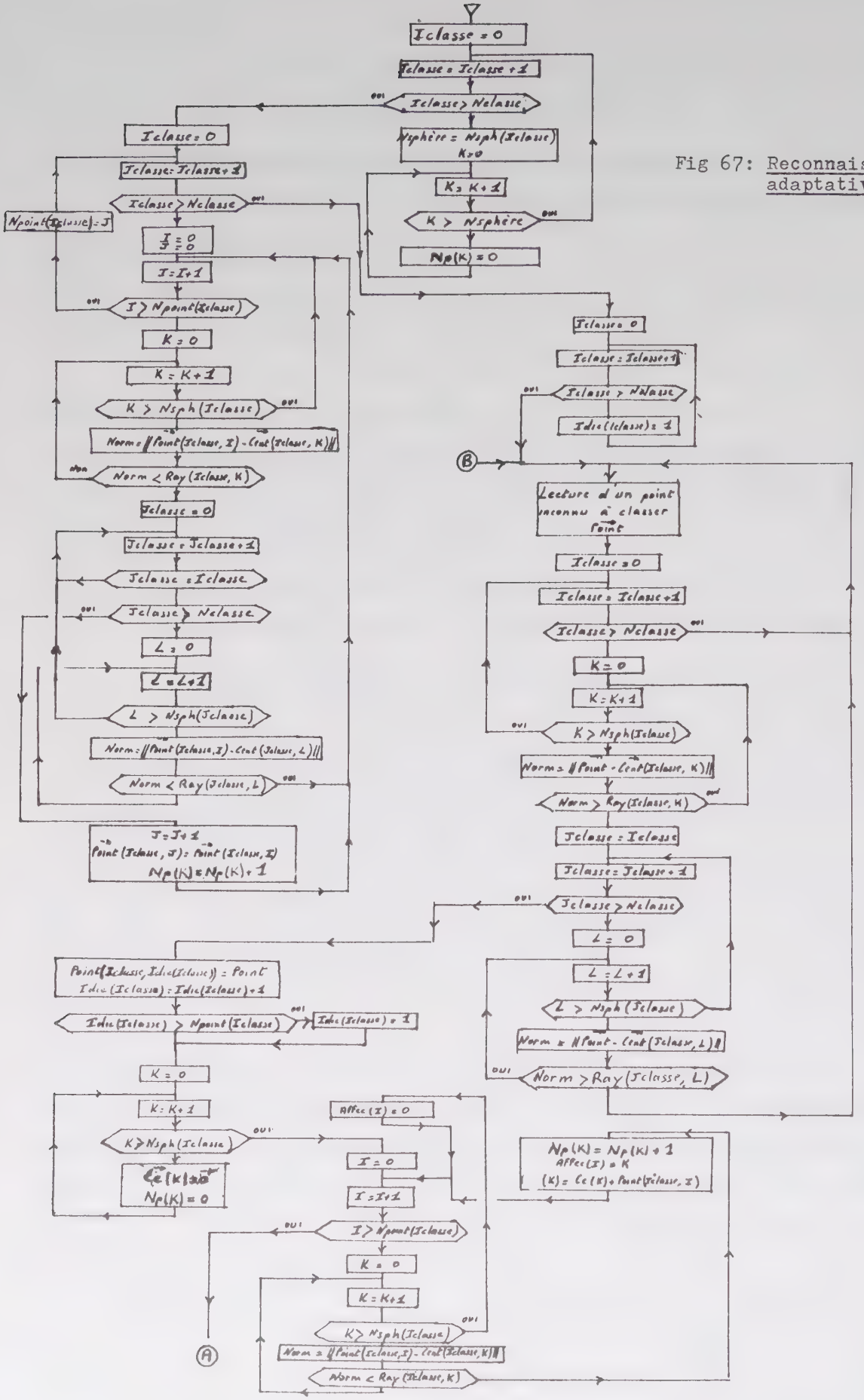


Fig 67: Reconnaissance adaptative

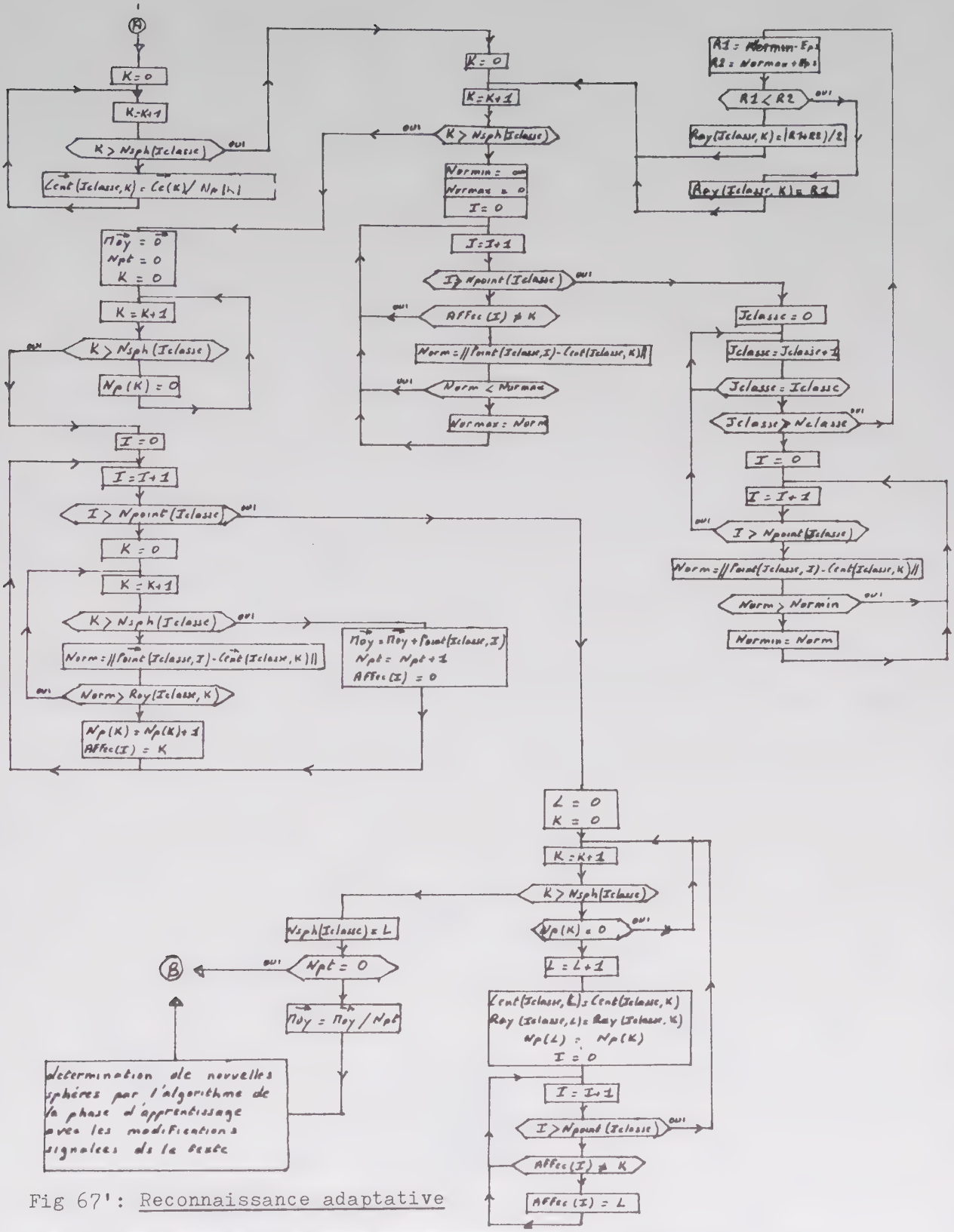


Fig 67': Reconnaissance adaptative

Liste des variables utilisées

Iclasse	: Indice de classe	Np(K)	: Nbr. pt. app. à la sphère K
Iclasse	: "	Npoint(I)	: Nbr. pt. app. à la classe I
Nclasse	: Nombre de classes	Nsph(I)	: Nbr. Sphère de la " I
K, L	: Indice de sphère	$\vec{Ce}$ (K)	: Centre des pts de la sphère K
I, J	: Indice de point	Affec(I)	: Sphère d'appartenance du pt I.
Norm	: Distance entre 2 points	Idis(I)	: Pointeur de la plus vieille forme de la classe I
$\vec{Point}$	: Point à classer	$\vec{Point}$(Iclasse, I)	: Coordonnées du pt I de la classe Iclasse
Nsphère	: Nombre de sphère	$\vec{Ray}$(I, J)	: Rayon de la sphère J de la classe I
Npt	: Nbr de pts non affectés	$\vec{Cent}$(J, I)	: Centre " " " " " " "
$\vec{Moy}$	: Centre en cours		

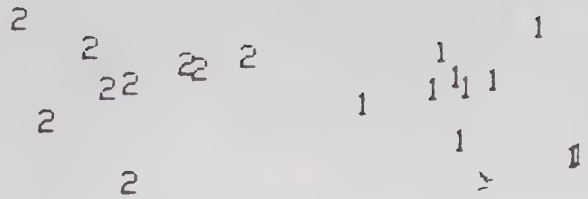


Fig 68

Fig 69

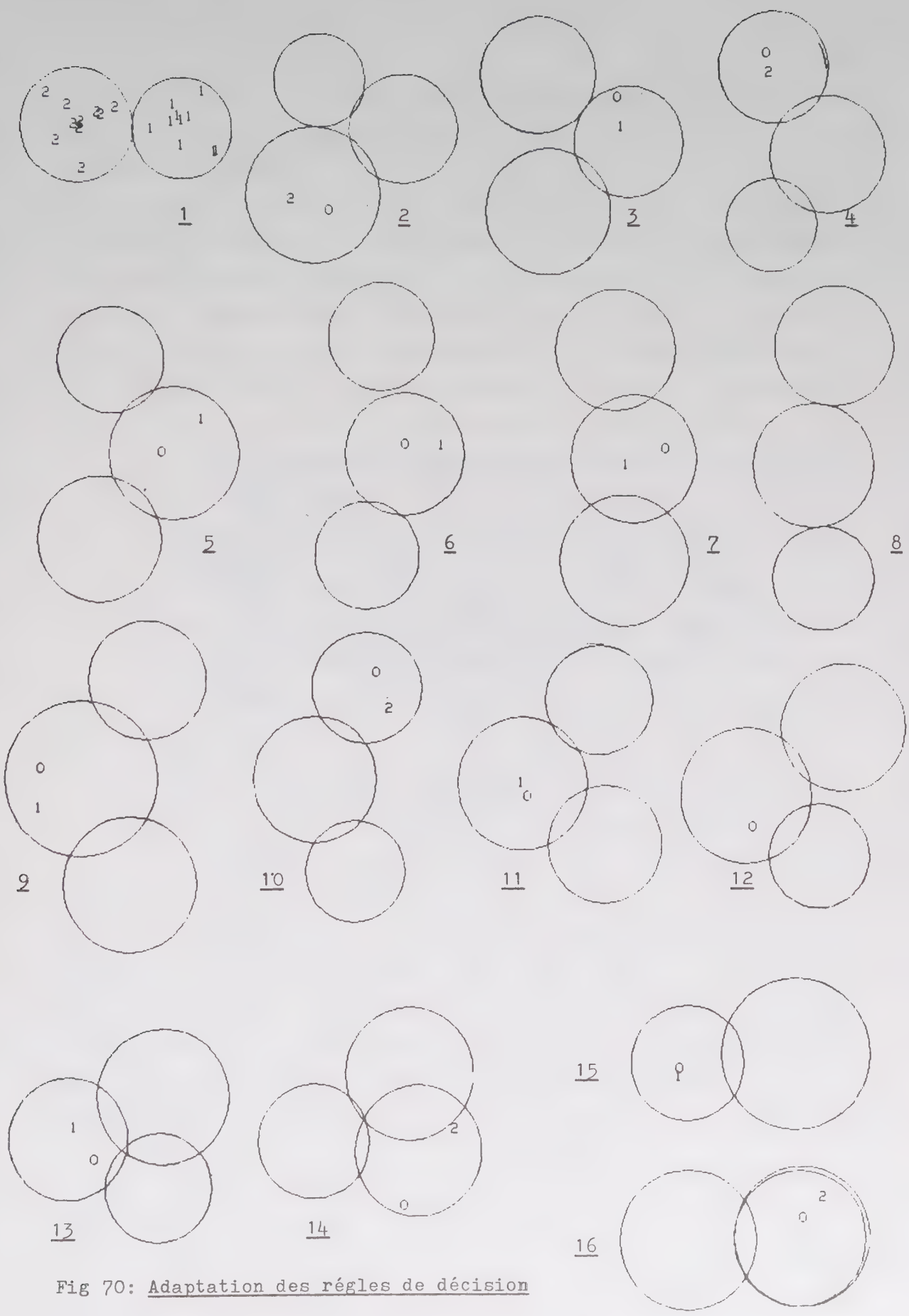


Fig 70: Adaptation des régles de décision

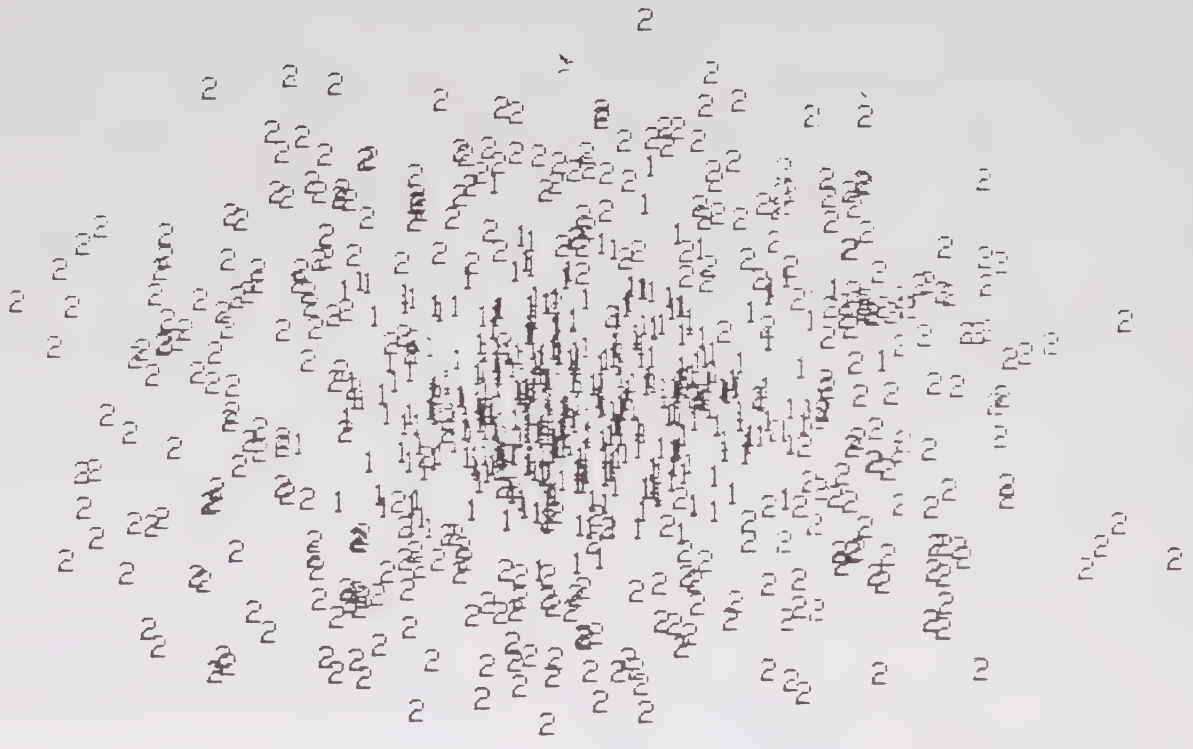


Fig 71

4.4.6. - Perspectives sur données nucléaires

L'algorithme de reconnaissance adaptative proposé ci-avant suppose que la forme représentée par un vecteur de $\mathbb{R}^n$ et que les dissemblances entre formes soient mesurées par une distance dérivant d'une norme. La méthode de paramètrisation utilisée au chapitre 3 ne peut donc être reconduite. La méthode utilisée par Dubuisson [27] qui consiste à découper les spectres en un certain nombre de bandes d'analyse nous semble convenir. Nous proposons le schéma d'application suivant :

- Découpe des signaux neutroniques en tranches de durée Tf.
- Calcul des spectres sur les huit voies.
- Moyennage bande de fréquence par bande de fréquence des énergies. La forme est alors représentée par un vecteur $\mathbb{R}^n$ $F = (x_1 -- x_n)$.
- Estimation des écarts types sur chaque coordonnée.
- Normalisation des formes de telle façon que la variance sur chaque coordonnée soit unitaire.
- Effectuer une classification automatique sur la base d'apprentissage (d'initialisation).

Nous sommes alors ramenés aux hypothèses d'utilisation de l'algorithme décrit aux paragraphes 4.4.2 et 4.4.4. De façon analogue au principe utilisé au chapitre 3 la décision d'anormalité serait prise lorsqu'une forme n'est pas jugée suffisamment fiable pour adapter les paramètres de l'algorithme.

5 - CONCLUSION

Ce mémoire expose une nouvelle application de la reconnaissance de formes au problème de la surveillance de processus complexes. Après une brève présentation du principe d'une chaîne de reconnaissance, nous proposons plusieurs façons de formaliser le problème suivant les données d'apprentissage disponibles et le but poursuivi (surveillance, contrôle, diagnostic). Nous exposons ensuite une méthode de surveillance du comportement vibratoire d'un coeur de réacteur nucléaire. Sont considérées comme formes : des tranches temporelles d'enregistrement simultanés, des signaux émis par les huit sections de détection neutronique du coeur. Les paramètres extraits de ces formes sont les caractéristiques des maxima des spectres des signaux les composant. L'apprentissage effectue une classification automatique sur une base de données normales. Lors de la phase de surveillance, la décision d'anormalité est prise lorsque la reconnaissance de la forme présentée en tant que membre de l'une des classes déterminées ne satisfait pas un critère de vraisemblance.

Aucune supposition n'est émise relativement à la nature des anormalités possibles. Le choix des algorithmes a été effectué sur des considérations générales et expérimentales en définissant l'anormalité par opposition à la base d'apprentissage supposée être un résumé exhaustif de ce qui doit être considéré comme normal. Bien que de nature heuristique les algorithmes utilisés ne nécessitent la définition que d'un petit nombre de paramètres. Ces paramètres sont déterminés soit de façon automatique, soit par des méthodes expérimentales explicitées. Deux critères servent à déterminer l'efficacité d'une méthode statistique : la probabilité de fausse alarme et la probabilité de détection. Une estimation de la probabilité de fausse alarme est effectuée en présentant à la phase de surveillance des données normales n'ayant pas servi à l'apprentissage.

La probabilité de détection est calculée sur deux types d'anormalités simulées : l'ajout d'une fréquence parasite et la présence d'un bruit blanc. Sur ces deux expériences les résultats obtenus montrent l'efficacité de la méthode de surveillance. Les algorithmes utilisés sont toutefois suffisamment simples et rapides pour être implantés en ligne sur micro-processeur.

Dans la deuxième partie de cette étude, nous évoquons le problème de l'évolution des critères de jugement et proposons un schéma de reconnaissance adaptative. Un algorithme basé sur une modélisation de chaque classe par des sphères est proposé et expérimenté sur des données simulées en dimension 2. Les résultats montrent qu'une bonne discrimination peut être obtenue dans des cas de non séparabilité linéaire. L'adaptativité de la reconnaissance est réalisée par un procédé permettant de faire varier en ligne les caractéristiques des modèles des classes (nombre de sphères, positions et rayons). Cette méthode a été testée sur des données simulées et assure des discriminations efficaces à chaque instant.

Une application sur données nucléaires n'a pu être réalisée toutefois nous avons montré comment les algorithmes proposés pourraient être utilisés.

L'approche du problème de la surveillance de processus par la reconnaissance de formes qui est de plus en plus utilisée dans les domaines de l'aéronautique, de la médecine et de la météorologie nous semble adaptée au domaine nucléaire. Des recherches sont effectuées dans ce sens tant en France qu'à l'étranger. Nous pensons que les années à venir verront l'utilisation effective de techniques de surveillance automatique basées sur des concepts d'intelligence artificielle.

REFERENCES

[1] DYDACK-SEGAL
Shape theory Springer 78

[2] PAVEL
Reconnaissance des structures dans les catégories : aspects algébriques et topologiques
Cours de 3ème cycle - Université Paris VI - 75/76

[3] PAU
Reconnaissance des formes et diagnostic des pannes
Cours de l'Ecole Nationale Supérieure des télécommunications polycopie 1978,

[4] G. BRACCHI, M. SOMALVICO
The correlation method for computed-aided statistical analysis, IEEE trans. Reliability, R 20, n°3, August 1971, 96-101

[5] L.D. FRENKEL, Application of computers to diagnostic medicine,
Southern Medical J., Septembre 1969, 1143-1148

[6] M.F. FRY
Importance des moyens de surveillance d'état des moteurs, l'Aéronautique et l'Astronautique, n° 42, 1973-4, 25-32

[7] D.B. DRANE, W.B. MARTIN, Pattern recognition applied to sleep state classification, Electroenceph. clin. Neurophysial., 1969, 26, p 238.

[8] J. PAGE, Recognition of patterns in jet engine vibrations signals,
IEEE Publ. 16 c 51, 102-105

[9] V.S. LEVADI, Automated learning applied to fault diagnosis, IEEE Trans. Aerospace and electronic systems, AES-3, n°6, Nov 1967, 941-946

[10] J.A. JONES,
The analysis of metallurgical data using pattern recognition techniques, Proc. Nat. Aerospace electronics conf., 17-19 - Mai 1971, IEEE Publ. 71-C-24-AES

[11] P.I. KUZNETZOV, L.A. PCHELINTZEV
The application of somme mathematical methods in medical diagnostics, Mathematical Biosciences, 5 (1969), 365-377

[12] R.L. HOFFMAN, K. FUKUNAGA,
Pattern recognition signal processing for mechanical diagnostics signature analysis, IEEE Trans. Computers, C-20, n° 9, Septembre 1969, 1095-1100

[13] W.J. HANKLEY, H.M. MERRILL,
A pattern recognition technique for system error analysis, IEEE Trans. Reliability, R-30, n° 3, August 1971, 148-153

[14] L.F. PAU
Application of pattern recognition to the diagnosis of equipment failures, Pattern Recognition Journal, 1974, Vol 6, n° 3-11.

[15] L.F. PAU,
Diagnosis of equipment failures by pattern recognition,
IEEE Trans. reliability, August 1974

[16] L.F. PAU
Diagnosis of equipment failures from dynamic measurements, Siemens foundation, Copenhague, Dec. 1973, 106 p

[17] L.F. PAU
Adaptive on line failure diagnosis and predictive pattern recognition, 2 nd Int. J. Conf. Pattern recognition, Copenhague, 13-15 août 1974, n° 6-9

[18] R.B. TATGE, Acoustical techniques for machinery diagnostics Acoustical Soc. of America Conf. Ottawa, 21-24 mai 1968

[19] L.F. PAU
Diagnostic statistique, Congrès National de fiabilité, Centre de fiabilité, CNET, PARIS, 20-22 septembre 1972 et, Onde électrique, Août-Septembre 1974, Vol 54, n°7

[20] C.Y. HSIUNG, C.W. COX,
Pattern classification in scan-type non-destructive tests, Int. J. of nondestructive testing. Vol 4 (1972) 231-247

[21] AL MOUHAMED
Methodes statistiques de détection d'anomalies de fonctionnement dans les processus complexes.
Thèse 3ème cycle - Université PARIS XI - 1977

[22] GONZALEZ - HOWINGTON
Reactor noise analysys by statistical pattern recognition methods.
Oak Ridge National Laboratory report 1977

[23] GONZALEZ - HOWINGTON - SIDES - KRYTER
A Multivariate Statiscal Pattern Recognition system for reactor noise analysis - IEEE Trans, On nuclear Sciences Vol, 23 N° 1 - 1976 - p. 342

[24] GONZALEZ - HOWINGTON
Dimensionality Reduction of reactor noise signatures
IEEE Trans on nuclear sciences and engineering
Vol. 62 - p. 163 - 1977

[25] GONZALEZ - FRY - KRYTER
Results in the application of pattern recognition methods to nuclear reactor core component surveillance,
IEEE - Nuclear Science 1974

[26] DUBUISSON - LAVISON
Utilisation des méthodes de reconnaissance des formes pour la surveillance d'un réacteur nucléaire
Revue générale nucléaire 1979 - N° 5 - Sept. Oct.

[27] DUBUISON - LAVISON
Surveillance of a nuclear reactor by use of a pattern recognition methodology
IEEE trans on systems, man and cylernetics
Vol SML 10 - N° 10 - Oct. 1980

[28] 13th informal meeting on reactor noise
Cadarache - May 7-9, 1980
Surveillance of a nuclear reactor by pattern recognition analysis of the neutronic noise

[29] BERNARD
Fluctuations neutroniques des réacteurs de puissance à eau sous pression
Thèse de 3ème cycle - Université PARIS XI - 1978

[30] IAEA - NPPCI - Specialist Meeting
Munich 5/7 décembre 1979
PWR Neutron noise surveillance : noise sources and their effects

[31] INVERNIZZI
Quelques méthodes et algorithmes utilisés en reconnaissance de formes
Note CEA N°

A N N E X E

REFERENCES BIBLIOGRAPHIQUES CLASSEES PAR THEMES

EXTRACTION DE PARAMETRES/ANALYSE DE DONNEES

- Mardia-Kent-Bibby : Multivariate analysis
  Academic Press - 1979

- Miclet-Grenier-Kaliszewski - Reconnaissance des formes
  Polycopie E.N.S.T - 1978

- Pau : Reconnaissance des formes et diagnostic des pannes
  Polycopie E.N.S.T - 1974

- Romeder : Méthodes et programmes d'analyse discriminante
  Dunod - 1973

- Sellam - Forcioli : Introduction de la notion d'écart entre sous espaces vectoriels en analyse de données
  RAIRO - Vol. 14 - N° 3 - Août 1980 - p. 283-301

- Giacomo della Ricia : Feature processing by optimal factor analyses techniques in statistical pattern recognition
  IEEE TRANS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE - VOL 2 - JANV. 1980

- Fernandez de la Vega : Détermination d'une fonction discriminante optimale.
  CONGRES AFCET-IRIA 1978 - Reconnaissance des formes et traitement des images.

EXTRACTION DE PARAMETRES/DEVELOPPEMENTS ORTHOGONAUX

- Watanabe : Evaluation and Selection of variables in pattern recognition. Computer and information sciences II - ed. by Tou Academic Press 1967.

- Tou-Heydorn : Some approaches to optimum feature extraction. Computer and information Sciences II - ed. by Tou Academic Press 1967.

- Chien-Fu : On the generalized Karhunen-Loève Expansion. IEEE Trans. on. information theory - July 1967.

- Tou-Gonzalez : Pattern recognition principles. Addison-Wesley Publishing Company 1974.

- Watanabe : Karhunen-Loeve and factor analysis Transactions of the 4th Prague conference on information theory 1965

- Christensen-Hirschman : Automatic phase alignment for the Karhunen-Loeve expansion. IEEE Trans. on biomedical engineering Vol. BME 26 - FEB. 1979

- Fukunaga-Koontz : Representation of random processes using the finite Karhunen- Loève expansion. Information and control - Vol. 16 - p. 85-101 - March 1970

- Jain : A sinusoidal family of unitary transforms IEEE trans. on pattern analysis and machine intelligence Vol. PAMI 1 - N° 4 - OCT. 1979.

- Kittler - Young : A new approach to feature selection based on the Karhunen-Loève expansion. Pattern recognition - Vol. 5 - p. 335-352 1973.

- Kittler : Features selection methods based on the Karhunen- Loève expansion. Pattern recognition theory and application. Ed. by. K.S. FU. Nato advanced study institues series, serie E N° 22 - 1977.

- Fukunaga-Koontz : Application of the Karhunen-Loève expansion to feature selection and Ordering. IEEE trans. on computers Vol. C19 nº 4 - April 1970.

- Patrick : Fundamentals of pattern recognition.
Prentice Hall information series 1972.

- Andrews : Multidimentional rotations in features selection
IEEE trans. on computers - Sept. 1971

- Kowalski - Bender : An orthogonal feature selection method.
Pattern recognition 1976 - Vol. 8 - p.p 1-4

EXTRACTION DE PARAMETRES/MESURES DE SEPARABILITE

- Mucciardi - Gose : A comparison of seven techniques for choasing subsets of pattern recognition properties.
  IEEE trans. on Computers C-20 - Sept. 71

- Michael - Lin ; Experimental study of information - Measure and inter-intra class distance ratios on features selection and ordering
  IEEE trans. on syst. man. Cybern SMC 3 Mars 73.

- Fang : A note on optimal selection of independant observables
  IEEE trans on syst. Man Cybern SMC9 - Mai 79

- Cover, Campenhout : On the possible orderings in the measurement selection problem.
  IEEE trans on syst. Man. Cybern SMC.7 - Sept. 77.

- Nelson - Levy : Selection of pattern features by mathematical programming algorithms.
  IEEE trans. on syst Science, Cybern SSC. 6 - Janv. 70

- Nelson - Levy : A dynamic programming approach to the selection of pattern features.
  IEEE trans. on. Syst Science, Cybern SSC.4 - Juillet 68

- Chang : Dynamic programming as applied to feature subset selection in a pattern recognition system.
  IEEE trans on Syst. Man. Cybern SMC.3 - MARS 73.

- Hussain - Donaldson : Suboptimal sequential decision schemes with on line feature ordering.
  IEEE trans on Computers - C.23 - Juin 74.

- Fu - Chien : A dynamic programming approach to sequential pattern Recognition.
  IEEE trans on electronic computers EC.16 - Déc. 67.

- Foley - Sammon : An optimal set of discriminant vectors
IEEE trans on computers C-24 - MARS 75.

- Edward Patrick : Fundamentals of pattern recognition
Prentice hall information series.

- Fukunaga - Short : A class of feature extraction criteria and its relation to the bayes risk estimate.
IEEE trans on inf. Theory IT 26 - Janv. 80.

- Peters : Feature selection for best mean square approximation of class densities.
Pattern. recognition - Vol.11 - p. 361-364 - 79.

- Duin - Van Haersma Buma - Roosma : On the evaluation of independent binary features.
IEEE trans on inf. Theory. IT 24 - Mars 78.

- Cover : The best two independent measurements are not the two best.
IEEE trans on syst. Man. Cybern. Janv. 74.

- Toussaint : Note on optimal selection of independent binary-valued features for pattern recognition.
IEEE trans. on Inf. Theory - Sept. 71.

- Yablon - Chu : The relationship of the bayes risk to certain separability measures in normal classification.
IEEE trans on pattern analysis and machine intelligence. PAMI2 - Mars 80.

- Boekee - Van der Luble : Some aspects of error bounds in feature selection.
Pattern recognition - Vol. 11 - p. 353-360 - 79.

- Devijver : On a new class of bounds on bayes risk in multihypothesis pattern recognition.
IEEE trans on computers - Vol. C23 - Janv. 74.

- Lissack - Fu : Error estimation in pattern recognition via L^{α} - Distance between posterior density functions.
IEEE trans on Inf. Theory IT 22 - Janv. 76.

- Toussaint : Some functional lower bounds on the expected divergence for multihypothesis pattern recognition, communication and radar systems.
IEEE trans on Syst. Man. Cybern. Oct. 71.

- Lianiotis : A class of upper bounds on probability of error for multihypothesis pattern recognition.
IEEE trans on inf. Theory - Nov. 69.

- Kailath : The divergence and battacharyya distance measures in signal selection.
IEEE trans on commun.Tech. COM-15 - 1967 - p. 52-60.

- Toussaint : An upper bound on the probability of misclassification in terms of the affinity.
Procceding of the IEEE - Fev. 77

- Toussaint : Comments on "The Divergenge and Bhattacha ryya distance measures in signal selection".
IEEE trans commun. COM-20 - Juin 72.

- Toussaint : Probability of error, Expected divergence, and the affinity of several distributions.

- Toussaint : Some properties of matusita's measure of afffinity of several distributions.
Ann. Inst. Stat. Math. Vol. 26 - N° 3 - 74.

- Chen : On the use of distance and information measures in pattern recognition and applications in pattern recognition theory and application,
Nato advanced studies - SERIE E - N° 22.

- Moshe Ben Bassat : On the sensitivity of the probability of error rule for feature selection.
IEEE trans on pattern analysis. Mach. Int. PAMI2 - Janv. 80.

- Hanakata : Feature selection and extraction for decision theoretic approach and structural approach. In pattern recognition theory and application.
Nato advanced studies SERIE E N° 22.

- Tou - Gonzalez : Pattern preprocessing and feature selection in pattern recognition principle chpt 7
Addison - Wesley Publishing Company - 1974

- Andrews : Feature selection
in : Introduction to Mathematical Techniques in Pattern Recognition
Chpt 2 - Wiley - Interscience.

- Narenda - Fukunaga : A branch and bound Algorithm for Feature Subset Selection.
IEEE trans on Computer - C.26 - Sept. 77.

- Kittler : Une généralisation de quelques algorithmes sous-optimaux de recherche d'ensembles d'attributs.
Congrès AFCET-IRIA - Recon. des formes et trait des images 78.

- Kittler : Feature Set Search Algorithms
in pattern recognition and signal processing
Nato advanced studies SERIE E N° 29.

- Moshe Ben Bassat : Myopic Policies in séquential classification
IEEE trans on computers C.27 - P 170-174 - 1978.

EXTRACTION DE PARAMETRES/TRANSFORMATIONS NON LINEAIRES

- Koantz - Fukunaga : A non linear feature extraction algorithm using distance transformation.
  IEEE Trans. On. Computers - Vol C.21 - Janv. 72

- Lee - SLAGLE - BLUM : A triangulation method for the sequential mapping of points from N space to 2 space.
  IEEE trans on computers - March 77.

- Bennet : The intrinsic dimensionality of signal collections.
  IEEE trans on information theory IT15 - 69.

- Chen - Andrews : Non linear intrinsic dimensionality computations.
  IEEE trans on computers - February 74

- Pettis - Bailey - Jain -Dubes : An intrinsic dimensionality estimator from near-neighbor information.
  IEEE trans on pat anal. and mach. int. PAMI1 - 79.

- Fukunaga - Olsen : An algorithm for finding intrinsic dimensionality of data.
  IEEE trans on computers - Vol. C20 - Feb. 71.

- Fukunaga - Short : Non linear feature extraction with a general criterion function.
  IEEE trans on inf. th. IT 24 - Sept.78

- Fukunaga - Ando : The optimum non linear features for a scatter criterion in discriminant analysis.
  IEEE trans on inf the. IT 23 - Juill. 77.

- Schartzman - Vidal : An algorithm for determining the topological dimensionality of point clusters.
  IEEE trans on comp. - Vol C.24 - Déc. 75

- Niemann - Weiss : A fast converging algorithm for non linear mapping of high dimensional data to a plane.
IEEE trans. on comp. Vol C28 - Feb. 79.

- Calvert : Nonorthogonal projections for feature extraction in pattern recognition.
IEEE trans on comp - May 70.

- Young : Classification estimation and pattern recognition chpt. 6 : feature extraction theory

APPRENTISSAGE

- Ball - Hall : A clustering technique for summarizing multivariate data behavioral sciences 12 - 153-155 (1967).

- From - Northouse : Class : a non parametric clustering algorithm. Patt. recogn. Vol. 8 - p 107 - 114.

- Delattre - Hansen : Bicriterion cluster analysis IEEE PAMI-2 - N° 4 [80] - p. 277.

- Gowda - Krishna : Aggloremative clustering using the concept of mutual nearest neighborhood Patt. Recogn. Vol. 10 - p. 105-112.

- DEVROYE - WAGNER : Asymptotic properties of clustering algorithms.

- Gitman - Levine : An algorithm for detecting unimodal fuzzy sets and its application as a clustering technique. IEEE C.19 - N7 (70) - p 583.

- Kittler : A comparative study of five locally sensitive clustering techniques. Internat. Comp. Sympos 1975.

- Leboucher - Lowitz : What a histogram can really tell the classifier. Patt. Recogn. Vol. 10 - p. 351 - 357.

- Gitman : An algorithm for non-supervised pattern classification. IEEE - SMC 3 (73) p. 66.

- Ruspini : New experimental results in fuzzy clustering. Information sciences - Vol. 6 - p 273 (73).

- Velasco - Rosenfeld : Some methods for the analysis of sharply bounded clusters. IEEE SMC-10 - p 511 (80).

- Jarvis - Patrick : Clustering using a similarity measure based on shared near-neighbors.
IEEE C-22 - p 1025 (73)

- Kittler : A locably sensitive method for cluster analysis.
Pattern recognition - Vol. 8 - p. 23 (76).

- Mizoguchi - Shimura : A non parametric algorithm for detecting clusters using hierachical structure.
IEEE PAMI-2 - p 292 (80).

- Koontz - Fukunaga : Asymptotic analysis of a non parametric clustering technique.
IEEE C-21 - p. 967 (72).

- Koontz - Fukunaga : A non parametric valley-seeking technique for cluster analysis.
IEEE C-21 - p 171 (72)

- Koontz-Narendra - Fukunaga : A graph-theoretic approach to non parametric cluster analysis.
IEEE - C.25 - p.936 (76).

- Bezdek : A convergence theorem for the fuzzy ISODATA clustering algorithms.
IEEE PAMI-2 - p.1 (80).

- Vogel-Wong : PFS clustering method
IEEE PAMI-1 - p 237 (79).

- Davies - Bouldin : A cluster separation measure
IEEE PAMI-1 - p 224 (79).

- Mardia - Kent - Bibly : Multivariate analysis
Academic Press 1979 - Chapt. 13 "Cluster Analysis".

- Everitt : Cluster analysis
Heinemann - Educational Books Ltd.

- Jambu : Classification automatique pour l'analyse des données.
Dunod (78).

- Diday : Optimisation en classification automatique.
INRIA.

- Kaufmann : Introduction à la théorie des sous-ensembles flous.
MASSON.

- Zahn : Graph-Theoretical methods for detecting and describing gestalt clusters.
IEEE C.20 - p 68 (71).

- Wright : Gravitational clustering
Pattern - recognition - Vol. 9 - p 151 (77).

DECISION

- Friedman : A recursive partitioning decision rule for non parametric classification.
  IEEE trans. on comp. 1977 - p 404-408.

- Dasarathy - Sheela : A composite classifier system design : Concepts and methodology.
  IEEE proccedings - Vol 67 - p 708-713.

- Beisner : A recursive baysian approach to pattern recognition.
  Pattern recognition 1968 - Vol. 1 - p 13-31.

- Chang : Pattern recognition by piecewise linear discriminant functions.
  IEEE trans. on Comp. C.22 - p 859-862.

- Fukunaga - Short : Generalized clustering for problem localisation.

- Sklansky - Michelotti : Locally trained piecewise linear classifiers.
  IEEE trans. on pattern analysis and machine intelligence. PAMI-2
  101 - 110.

- Wassel - Sklansky : An adaptative non parametric linear classifier.
  Proccedings of IEEE - Vol. 64 - 1976 - p 1162 - 1171.

- Perennou : Algorithme pour l'optimisation d'un discriminateur.
  Laboratoire CERFIA - Université de Toulouse.

- Chang : Finding prototypes for nearest neighbor classifiers.
  IEEE trans on comp. C.23 - p 1179 - 1184.

- Delannoy : Un algorithme de recherche de plus proches voisins.
  RAIRO - Vol. 14 - N° 3 - 1980 - p 275 - 286.

- Dill : Quelques extensions de la méthode des plus proches voisins,
  Congrès AFCET - IRIA "Reconnaissance formes" - 1979 - p 292 - 299.

- Devijver : Décision-theoretic and related approaches to pattern classification.
in : "Pattern recognition theory and applications"
Nato advanced sludies - SERIE E - N° 22.

- DEVROYE : On the inequality of Cover and Hart in nearest neighbor discrimination.
IEEE PAMI-3 - 1981 - p 75 - 78.

- Patrick - Fisher : A generalized k-nearest neighbor rule .
Information and control 16 - p 128 - 152.

- Cover : Geometrical and statistical properties of systems of linear inequalities with application in pattern recognition.
IEEE trans. on Electronic Computers EC.14 - p 326.

- Cooper : Adaptative pattern recognition and signal detection using stochastic approximation.
IEEE trans. on electronic computers EC. 13 - p. 306 - 307.

- Chow : On optimum recognition error and reject trade off.
IEEE. trans. on. Information theory - IT 16 - p 41 - 46.

- Hellman : The nearest neighbor classification rule with a reject option.
IEEE trans. on systems. science cylernetics SSE6 - p 179 - 185.

- Hart : The condensed nearet neigbor rule.
IEEE trans on information theory - IT 14 - p 515.

- HO-Kashyap : An algorithm for linear inequalities and its applications.
IEEE trans on electronic computer EC 14 - p 683 - 688.

- Friedman - Baskett - Shustek : An algorithm for finding nearest neighbors.
IEEE trans on computer C.25 - p 1000 - 1006.

- Dasarathy - white ; A characterization of nearest - neighbor rule decision surfaces and a new approach to generate them.
Pattern recognition - Vol. 10 - p 41 - 46.

- Duda - Fossum : Pattern classification by iteratively determined linear and piecewise linear discriminant fonctions.
IEEE trans on electronic computer EC 15 - p 220 - 231.

Manuscrit reçu le 10 juin 1982